

Determinanten und Prüfung der Nachhaltigkeitsberichterstattung

Dissertation
eingereicht an der Wirtschaftswissenschaftlichen Fakultät
der Universität Zürich

zur Erlangung der Würde
eines Doktors
der Wirtschaftswissenschaften, Dr. oec.

vorgelegt von

Christian Schlick
von Deutschland

genehmigt im April 2015 auf Antrag von

Prof. Dr. Dieter Pfaff
Prof. Dr. Conrad Meyer

Die Wirtschaftswissenschaftliche Fakultät der Universität Zürich gestattet hierdurch die Drucklegung der vorliegenden Dissertation, ohne zu den darin ausgesprochenen Anschauungen Stellung zu nehmen.

Zürich, 01.04.2015

Der Vorsitzende des Doktoratsausschusses: Prof. Dr. Josef Zweimüller

Meinem Vater

Dank

Es ist mir ein grosses Anliegen, allen Personen zu danken, die zum Gelingen dieser Dissertation beigetragen haben. Meinem Doktorvater Prof. Dr. Dieter Pfaff danke ich für den Freiraum bei der thematischen Auswahl und Gestaltung der Papers, die wertvolle Unterstützung im Entstehungsprozess sowie die Erkenntnis, dass sich Wahrheit im Dialog ereignet. Herrn Prof. Dr. Conrad Meyer gebührt mein Dank für die hilfreichen Hinweise während der internen Doktorandenseminare und die Übernahme des Zweitgutachtens. Besonderer Dank gilt meiner Kollegin Frau Dr. Katrin Hummel. Es hat mir grosse Freude bereitet, mit ihr bestehende Ansätze zu hinterfragen, neue Sichtweisen zu entwickeln und an der Klärung offener Fragen im Forschungsgebiet der Nachhaltigkeitsberichterstattung und deren Prüfung zu arbeiten. Ich danke auch meinen Kolleginnen und Kollegen Herrn Dr. Sandro Kürschner, Herrn Dr. Peter Ising, Frau Patricia Ruffing, Frau Silvia Chibane, Frau Dr. Ruta Bilkeviciute, Herrn Dr. André Meier, Herrn Benedikt Bisig und Herrn Stefan Beeler für die vielen Anmerkungen und ihr offenes Ohr in allen Phasen der Dissertation. Darüber hinaus möchte ich meinen Kolleginnen und Kollegen am Lehrstuhl von Herrn Prof. Dr. Conrad Meyer für den offenen Gedankenaustausch danken. Ausserdem gilt mein Dank den Teilnehmern der Konferenzen und Seminare sowie den Gutachtern im Zulassungs- und Veröffentlichungsprozess für ihre Kommentare zur Weiterentwicklung der Papers.

Besonders möchte ich meiner Partnerin Franziska Brüning, meinen Eltern Dres. Heike und Wolf-Rüdiger Schlick sowie meinem Bruder Max Schlick für ihre Unterstützung und Geduld danken.

Zürich, im April 2015

Christian Schlick

Inhaltsverzeichnis

1. Thematischer Rahmen	8
1.1 Unternehmensberichterstattung und Nachhaltigkeit	8
1.2 Forschungsgebiete	9
1.3 Beitrag und Zusammenfassung der eingebrachten Papers	14
Literaturverzeichnis	16
2. Eingebrachte Papers	22
Zusammenhang zwischen Nachhaltigkeitsperformance und Nachhaltigkeitsbericht- erstattung – Legitimität oder finanzielle Überlegungen?	22
The Relationship between Sustainability Performance and Sustainability Disclosure Quality – Hard Numbers Beat Smooth Talk	48
The Nature of Managerial Capture in Sustainability Assurance Engagements.....	80

1. Thematischer Rahmen

1.1 Unternehmensberichterstattung und Nachhaltigkeit

Die externe Berichterstattung von Unternehmen unterliegt einem stetigen Wandel. Dieser betrifft zum einen die klassische Finanzberichterstattung, in der eine zunehmende Regeldichte zu neuen und geänderten Offenlegungspflichten beiträgt. Dabei wirken internationale Konvergenzbestrebungen dem heterogenen Bild nationaler Rechnungslegungsstandards entgegen (FASB, 2014).¹ Gleichzeitig werden die Unternehmen mit sich wandelnden nationalen handelsrechtlichen Vorschriften², börsenspezifischen Kotierungsreglementen³ sowie länderspezifischen Richtlinien etwa zur Corporate Governance⁴ konfrontiert. Zum anderen führt die zunehmend internationale Tätigkeit von Unternehmen zur Globalisierung von aus Unternehmenssicht externen Effekten, die sich auf eine Vielzahl von Gesellschaftsgruppen sowohl positiv als auch negativ auswirken können (Coase, 1960; Pigou, 1932). Das Bewusstsein über die komplexe Problematik solcher Wechselwirkungen induzierte einen internationalen Reflexionsprozess darüber, ob und inwieweit zum Gesamtwohl der Gesellschaft zusätzliche Anstrengungen zur stärkeren Berücksichtigung von Externalitäten in unternehmerischen Entscheidungsprozessen notwendig sind. Mit der weit gefassten Definition von nachhaltiger Entwicklung als “development [that] seeks to meet the needs and aspirations of the present without compromising the ability to meet those of the future” (United Nations, 1987, Teil I, Kapitel 1, Absatz 49) wird deutlich, dass eine nachhaltige Unternehmensführung kurz- und langfristige externe Effekte umfassend einbeziehen und wo nötig steuern muss. Aus diesem übergeordneten Zusammenhang leitet sich für die Berichterstattung der Unternehmen ab, dass ausgehend von der klassischen Finanzberichterstattung eine Erweiterung notwendig ist. Es sind zusätzliche Informationen offenzulegen, die den individuellen Beitrag eines Unternehmens zu einer breit verstandenen Nachhaltigkeit aufzeigen. So können nicht nur die Informationsbedürfnisse der von externen Effekten betroffenen Interessengruppen befriedigt werden, sondern es entstehen gleichzeitig Anreize für die Unternehmen, sich mit dem Thema Nachhaltigkeit neu auseinanderzusetzen. Als Ergebnis wurden den Stakeholdern⁵ von Unternehmen erstmals in den 1970er Jahren speziell aufgearbeitete Informationen über die sozialen und umweltbezogenen Auswirkungen der Unternehmensaktivitäten bereitgestellt (Fifka, 2012). Nach einer Entwicklungszeit von rund 40 Jahren ist heute zumindest für internationale Grossunternehmen eine umfassende Nachhaltigkeitsberichterstattung “undeniably a mainstream global business practice” (KPMG, 2013, S. 22).

¹ Vgl. zum Stand der nationalen IFRS-Adaption IASB (2014), PricewaterhouseCoopers (2014) und Deloitte (2014).

² Vgl. zum Beispiel das Deutsche Bilanzrechtsmodernisierungsgesetz vom 25. Mai 2009 oder das revidierte Schweizer Rechnungslegungsrecht, welches am 1. Januar 2013 in Kraft getreten ist.

³ Vgl. zum Beispiel für Deutschland Deutsche Börse (2014) und für die Schweiz SIX Swiss Exchange (2014).

⁴ Vgl. die Comply-or-explain-Richtlinien des Deutschen Corporate Governance Kodex in der Fassung vom 24. Juni 2014 in Verbindung mit § 161 AktG sowie den Swiss Code of Best Practice for Corporate Governance in der Fassung vom 28. August 2014.

⁵ Freeman (1984, S. 25) definiert Stakeholder als “groups [...] that can affect, or are affected by, the accomplishment of organizational purpose”.

1.2 Forschungsgebiete

Parallel zu den Entwicklungen in der Unternehmenspraxis hat sich rund um die Nachhaltigkeitsberichterstattung ein Fundus an Forschungsarbeiten entwickelt, der sich in drei Forschungsgebiete untergliedern lässt. Das Forschungsgebiet I umfasst Arbeiten, die sich mit der Operationalisierung von Nachhaltigkeit beschäftigen. Es geht um die Fragen, was Nachhaltigkeit im Unternehmenskontext konkret bedeutet und anhand welcher Indikatoren die Nachhaltigkeit eines Unternehmens beurteilt werden kann. Im Report der Vereinten Nationen von 1987 wird neben der Definition von nachhaltiger Entwicklung auch die Bedeutung der Wirtschaft und somit einzelner Unternehmen bei der Entstehung und Lösung globaler Probleme, wie zum Beispiel dem Klimawandel, der Energie- und Nahrungsversorgung, Ressourcenknappheit, Gesundheit, Bildung und dem Schutz von Ökosystemen, beschrieben (United Nations, 1987). Auf Basis dieser breiten Grundlage werden verschiedene theoretische Nachhaltigkeitskonzepte diskutiert (Marshall & Toffel, 2005), wobei die breiteste Akzeptanz dem Triple-Bottom-Line-Ansatz zugesprochen werden kann (Marquis & Toffel, 2011). Nach diesem Konzept umfasst unternehmerische Nachhaltigkeit die drei Bereiche Ökonomie, Ökologie und Soziales (Elkington, 1997). Bei der Gewichtung dieser Bereiche kann zwischen einem schwachen und starken Nachhaltigkeitsverständnis unterschieden werden, je nach dem ob die Summe aller drei substituierbaren Bereiche mindestens konstant bleibt (schwach) oder ob dabei zusätzlich nur regenerative ökologische Ressourcen verbraucht werden (stark) (Ott & Döring, 2008). Da Unternehmen ohne eine stabile ökonomische Grundlage auf Dauer ihrer ökologischen und sozialen Verantwortung nicht nachkommen können, sehen Kritiker des Triple-Bottom-Line-Ansatzes die Gefahr, dass einzelne Unternehmen in der Triple-Bottom-Line eine “good old-fashioned single bottom line plus vague commitments to social and environmental concerns” (Norman & MacDonald, 2004, S. 256) sehen könnten. Aus Sicht der Unternehmen wandelt sich die Bedeutung der drei Dimensionen zunehmend in Richtung eines interdependenten und integrativen Verständnisses. Während in den 1970er und 1980er Jahren der Terminus Sozialbericht verbreitet war, dominierten in den 1990er Jahren Umweltberichte und in den 2000er Jahren die Termini (sozialer) Verantwortungsbericht und Nachhaltigkeitsbericht (Fifka, 2012). In jüngster Zeit werden Nachhaltigkeitsberichte vermehrt anstelle einer separaten Veröffentlichung in den Jahresbericht integriert, was die Zunahme der ihnen zugemessenen Bedeutung unterstreicht (KPMG, 2013).

Das theoretisch und praktisch weiterhin uneinheitliche Verständnis von unternehmerischer Nachhaltigkeit setzt sich bei der Messung der Nachhaltigkeitsperformance und der Beurteilung von Quantität sowie Qualität der Nachhaltigkeitsberichterstattung fort. Die ökonomische Dimension der Nachhaltigkeitsperformance lässt sich auch als finanzielle Unternehmensperformance interpretieren (Wood, 2010), zu deren Messung sich beispielsweise Aktien- und Eigenkapitalrenditen oder Ergebnisgrößen verwenden lassen (Orlitzky et al., 2003; Ullmann, 1985). Während sich solche Masse auf direkt beobachtbare Resultate beziehen, wird im Bereich der sozialen und ökologischen Dimension auch die Berücksichtigung indirekt beobachtbarer Anstrengungen der Unternehmen diskutiert. Hierunter fallen die Prinzipien der Unternehmensführung und die Ausgestaltung von nachhaltigkeitsbezogenen internen Prozessen (Wood, 1991). Die in empirischen Studien verwendeten Messkonzepte sind breit gefächert

und decken je nach Studie direkte, indirekte oder gemischte Performancemasse ab, wobei die benötigten Daten unter anderem aus den Nachhaltigkeitsberichten der Unternehmen erhoben werden können. Beispiele für direkte Masse sind normalisierte Mengen oder Recyclingraten giftiger Abfälle und giftiger Emissionen generell (Aerts & Cormier, 2009; Al-Tuwaijri et al., 2004; Clarkson et al., 2008; Clarkson et al., 2011; Patten, 2002). Ausserdem kommen Rankings zum Einsatz, die mehrere direkte Umweltparameter aggregieren (Cho et al., 2012). Als gemischte Performancemasse werden weitere Rankings verwendet, die sowohl direkte Emissionswerte als auch teilweise Prinzipien und Prozesse sowie nachhaltigkeitsbezogene Stärken und Schwächen berücksichtigen (Ingram & Frazier, 1980; Wiseman, 1982; Rockness, 1985; Fekrat et al., 1996; Hughes et al., 2001; Cho et al., 2006; Cho & Patten, 2007; de Villiers & van Staden, 2011). Weiterhin kommen die Industriegruppe und die Meldung von Emissionen an öffentliche Stellen als Proxies zum Einsatz (Bewley & Li, 2000).

Bei der Messung der Nachhaltigkeitsberichterstattung stehen die Fragen im Mittelpunkt, wie und im Vergleich zu welcher Benchmark die Quantität und Qualität der veröffentlichten Informationen bewertet werden können. Bei den berichtenden Unternehmen haben sich die Leitlinien zur Berichterstattung der Global Reporting Initiative (GRI) durchgesetzt (Hoffmann, 2011; KPMG, 2013). Gemäss GRI umfasst die Nachhaltigkeitsberichterstattung “the practice of measuring, disclosing, and being accountable to internal and external stakeholders for organizational performance toward the goal of sustainable development” (GRI, 2011, S. 3). Die Leitlinien umfassen nachhaltigkeitsbezogene Prinzipien, Prozesse sowie Performancekennzahlen und berücksichtigen ausserdem branchenspezifische Besonderheiten in der Relevanz verschiedener Indikatoren (GRI, 2013a, 2013b). Daneben existieren unternehmensinterne Richtlinien sowie weitere Empfehlungen von Branchenverbänden (Verband der Chemischen Industrie et al., 2013; International Council on Mining and Metals, 2014). In Europa ist die gesetzliche Regelungsdichte zur Nachhaltigkeitsberichterstattung verglichen mit der klassischen finanziellen Berichterstattung gering, nimmt jedoch tendenziell zu.⁶ Parallel zur umfangreicher und detaillierter werdenden Berichterstattung in der Praxis kommen in Forschungsarbeiten zunehmend komplexere Benchmarks zur Beurteilung der Nachhaltigkeitsberichterstattung zum Einsatz, so zum Beispiel Ansätze in Anlehnung an die GRI-Leitlinien (Cho et al., 2012; Clarkson et al., 2008; Clarkson et al., 2011). Sie ersetzen einfachere Methoden der Inhaltsanalyse, die auf eine geringere Anzahl von teilweise nach Informationstyp gewichteten Berichtselementen zurückgreifen (de Villiers & van Staden, 2011; Aerts & Cormier, 2009; Patten, 2002; Cho & Patten, 2007; Cho et al., 2006; de Villiers & van Staden, 2006; Al-Tuwaijri et al., 2004; Hughes et al., 2001). Daneben kommen auch indirekte Messkonzepte zum Einsatz, die zum Beispiel auf der Reaktion von Unternehmen auf Umfragen oder der Wahrnehmung der Berichterstattung durch die Teilnehmer von Experimenten beruhen (Dawkins & Fraas, 2011; Rockness, 1985).

⁶ Vgl. zum Beispiel die verpflichtenden Angaben zu Kohlendioxidemissionen in Grossbritannien ab 2013 (Companies Act 2006 Regulations 2013). Ausserdem müssen Schwedische Unternehmen im Staatsbesitz gemäss den GRI-Leitlinien berichten (Richtlinien zur externen Berichterstattung für staatliche Unternehmen, 2007). In Frankreich sind seit 2012 zusätzliche Berichtspflichten wirksam (Grenelle II, 2010). Der Europäischen Kommission wurde am 16. April 2013 die Gesetzgebungsvorlage C7-0103/13 zur Offenlegung nichtfinanzieller Informationen vorgelegt, die nach der Umsetzung in nationales Recht zu zusätzlichen Berichtspflichten führen kann.

Im Forschungsgebiet II wird Fragen nach dem Nutzen und den Determinanten unternehmerischer Nachhaltigkeit aus der Perspektive der Unternehmen nachgegangen. Von besonderem Forschungsinteresse ist die Beziehung zwischen der ökonomischen Dimension der Nachhaltigkeit und ihren beiden anderen Dimensionen, wobei “attempts to relate social [and environmental] performance to economic performance were undertaken largely in the hopes of establishing a positive relationship that might be persuasive to business leaders who were skeptical about the value” (Wood, 2010, S. 60). Aus theoretischer Sicht konkurrieren hier verschiedene Ansätze zum Umgang mit umfassender unternehmerischer Nachhaltigkeit. Nach dem traditionellen Ansatz optimiert allein das eher kurzfristige Streben der Unternehmen nach Gewinnmaximierung die Gesamtwohlfahrt, während der aufgeklärte Ansatz langfristige unternehmerische Eigeninteressen berücksichtigt und zusätzlichen Zielgrößen eine Bedeutung zuspricht. Hinzu kommt der verantwortungsbewusste Ansatz, nach dem sich umfassende Nachhaltigkeit für das einzelne Unternehmen möglicherweise nicht auszahlt, sie aber dennoch der gesamtgesellschaftlich wünschenswerte Weg ist. Ausserdem könnten Unternehmen nach dem konfusen Ansatz ihre Nachhaltigkeitsbestrebungen mit ethischen Argumenten rechtfertigen, aber trotzdem einen finanziellen Nutzen erwarten (McGuire, 1969). Die empirischen Ergebnisse zum Zusammenhang zwischen finanzieller Unternehmensperformance und den beiden verbleibenden Dimensionen der Nachhaltigkeit deuten in Meta-Analysen trotz aller methodischen Schwierigkeiten insgesamt und über verschiedene Industriegruppen hinweg auf einen positiven Zusammenhang hin (Orlitzky et al., 2003). “After 35 years of research and over 200 studies, there is a conclusive [...] answer to the question of whether companies benefit financially from social performance. The effect of CSP [corporate social performance] on CFP [corporate financial performance] is small, positive, and significant” (Margolis et al., 2009, S. 28).⁷

In der Frage nach den Determinanten der Nachhaltigkeitsberichterstattung besteht bislang weder hinsichtlich des theoretischen Hintergrunds noch der empirischen Forschungsergebnisse ein einheitliches Bild. Einigkeit besteht jedoch weitgehend darin, dass aus Sicht der Unternehmen bei der Auswahl und Darstellungsweise konkreter Berichtsinhalte nicht nur die Informationsbedürfnisse der Stakeholder, sondern besonders unternehmenseigene strategische Interessen von Bedeutung sind (Cho et al., 2012; Clarkson et al., 2011; Clarkson et al., 2008; Cho & Patten, 2007; de Villiers & van Staden, 2006; Al-Tuwaijri et al., 2004; Patten, 2002). Die Vertreter der Theorie der freiwilligen Offenlegung gehen davon aus, dass ein Unternehmen private Informationen solange zurückhält, bis die erwarteten positiven Effekte auf den Marktwert des Unternehmens die erwarteten negativen Effekte auf ebendiesen übersteigen (Verrecchia, 1983). Übertragen auf die Nachhaltigkeitsperformance eines Unternehmens folgt daraus, dass umso mehr Nachhaltigkeitsinformationen veröffentlicht werden, je besser die Performance ist – es besteht demnach ein positiver Zusammenhang zwischen beiden Größen (Clarkson et al., 2008). Im Gegensatz dazu gehen Vertreter der Legitimitätstheorie von einem negativen Zusammenhang aus. Als die Legitimität eines Unternehmens wird “a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions” (Suchman, 1995, S. 574) verstanden. Sollte die Nachhaltigkeitsperformance eines Unternehmens den

⁷ CSP umfasst die ökologische und die soziale Dimension der Nachhaltigkeit.

gesellschaftlichen Normen widersprechen, so kann ein Unternehmen seine Legitimität verlieren (Dowling & Pfeffer, 1975) und zentrale Stakeholder können ihre für den Unternehmenserfolg wesentliche Unterstützung verweigern (Davis, 1973; Wood, 1991). Ein Beispiel dafür ist ein unvorteilhaftes Unternehmensimage, welches zu unzufriedenen Kunden, Umsatzrückgang, Schwierigkeiten bei der Mitarbeitergewinnung und -bindung sowie einem höheren Risiko von Rechtsstreitigkeiten und regulatorischen Auflagen führen kann (Ameer & Othman, 2012; Ihlen & Roper, 2014; Kollman & Prakash, 2001; Wood, 1991). “Legitimacy is mainly about perception” (Aerts & Cormier, 2009, S. 1), das heisst, eine Entkopplung der Wahrnehmung eines Unternehmens durch seine Stakeholder von seinem tatsächlichen Zustand der Nachhaltigkeitsperformance ist möglich. Daraus lassen sich verschiedene Strategien zur Bewahrung der Legitimität ableiten. Es wird empfohlen, zuerst die öffentliche Wahrnehmung durch die Offenlegung von Informationen zur Nachhaltigkeit zu beeinflussen, danach Handlungen mit Symbolcharakter durchzuführen und nur in letzter Instanz die Nachhaltigkeitsperformance selbst zu ändern (Sethi, 1978). Neuere Ansätze unterscheiden zwischen Strategien zur Vermeidung von Legitimitätsrisiken, zur Änderung gesellschaftlicher Normen, zur Beeinflussung öffentlicher Wahrnehmungen und zur Erreichung einer Übereinstimmung mit gesellschaftlichen Erwartungen (O’Donovan, 2002). In diesem Rahmen ist auch eine gezielte Beeinflussung der Stakeholder durch die Unternehmensberichterstattung möglich, wodurch die Nachhaltigkeitsberichterstattung zu einem Werkzeug zur Sicherstellung der Legitimität eines Unternehmens wird (Deegan, 2002). So können Unternehmen beispielsweise Greenwashing betreiben, indem sie von kritischen Themen durch eine selektive und besonders umfangreiche Veröffentlichung positiver Informationen ablenken (Lyon & Maxwell, 2011). In diesem Fall besteht zwischen der tatsächlichen Nachhaltigkeitsperformance und der Quantität der Nachhaltigkeitsberichterstattung ein negativer Zusammenhang.

Vor dem Hintergrund der gegensätzlich interpretierten Implikationen von freiwilliger Offenlegungstheorie und Legitimitätstheorie kommen empirische Studien beim Einfluss der Nachhaltigkeitsperformance auf die Quantität der Berichterstattung zu positiven Ergebnissen (Al-Tuwaijri et al., 2004; Bewley & Li, 2000; Clarkson et al., 2008), negativen Ergebnissen (Cho et al., 2012; Clarkson et al., 2011; Dawkins & Fraas, 2011; Cho & Patten, 2007; Cho et al., 2006; de Villiers & van Staden, 2006; Patten, 2002; Hughes et al., 2001) oder können keinen Zusammenhang nachweisen (Fekrat et al., 1996; Ingram & Frazier, 1980; Rockness, 1985; Wiseman, 1982). Neben der Nachhaltigkeitsperformance selbst und der Nachhaltigkeitsstrategie eines Unternehmens wird auch von einem Einfluss der Unternehmensgrösse (Branco & Rodrigues, 2008; Clarkson et al., 2008; Clarkson et al., 2011; Cormier et al., 2005; Dawkins & Fraas, 2011; Kolk, 2003; Patten, 2002), der finanziellen Performance (Cormier & Magnan, 2003; Cormier et al., 2005; Neu et al., 1998; Ullmann, 1985) sowie des Verschuldungsgrads (Branco & Rodrigues, 2008; Clarkson et al., 2008; Cormier & Magnan, 2003) ausgegangen. Es konnten weiterhin länderspezifische Unterschiede vor dem Hintergrund abweichender regulatorischer Rahmenbedingungen und kultureller Varianzen (Fekrat et al., 1996; Kolk, 2003; Kolk et al., 2001; Orij, 2010; van der Laan Smith et al., 2005) sowie industriespezifische Unterschiede (Cho & Patten, 2007; Dawkins & Fraas, 2011; Patten, 2002; Roberts, 1992) gezeigt werden.

Das Forschungsgebiet III beschäftigt sich mit der Prüfung von Nachhaltigkeitsberichten. Hier werden die Determinanten der Wahlentscheidung zur externen Prüfung von Nachhaltigkeitsberichten untersucht sowie die Aussagekraft der veröffentlichten Prüfberichte und Aspekte der Prüfungsdurchführung. Die Entscheidung zur Prüfung des Nachhaltigkeitsberichts erfolgt in Europa weitgehend freiwillig, jedoch ist auch hier eine zunehmende Regelungsdichte zu beobachten.⁸ Empirische Studien kommen zum Ergebnis, dass Prüfungen häufiger in Ländern mit breiter Stakeholder-Orientierung und somit einer relativ geringeren Bedeutung der Shareholder durchgeführt werden und ausserdem institutionelle Faktoren von Bedeutung sind (Kolk & Perego, 2010; Simnett et al., 2009). Weiterhin beeinflusst die Industriezugehörigkeit sowie die Unternehmensgrösse die Entscheidung zur Prüfung (Cho et al., 2014; Simnett et al., 2009). Die Aussagekraft der veröffentlichten Prüfberichte steht in der Kritik, was auf die fehlende Einbindung externer Stakeholder in den Prüfprozess und damit den massgeblichen Einfluss von Eigeninteressen der Unternehmen zurückgeführt wird (Ball et al., 2000; Cooper & Owen, 2007; Deegan et al., 2006; O'Dwyer & Owen, 2005). Der zugehörige Begriff des Managerial Capture "sees management take control of the whole process [...] by strategically collecting and disseminating only the information it deems appropriate to advance the corporate image, rather than being truly transparent and accountable to the society it serves" (Owen et al., 2000, S. 85). Auf der Ebene der Prüfungsdurchführung wird deutlich, dass zwischen Prüfer und geprüftem Unternehmen ein dichtes und iteratives Beziehungsgeflecht besteht (O'Dwyer, 2011; O'Dwyer et al., 2011; Park & Brorson, 2005). Neben der negativen Interpretation dieses Geflechts vor dem Hintergrund von Managerial Capture liegen auch Ergebnisse vor, welche die Unabhängigkeit des Prüfers hervorheben und nach denen sich Prüfungsgesellschaften zu ihrer Verantwortung gegenüber externen Stakeholdern bekennen (O'Dwyer, 2011; O'Dwyer et al., 2011). Hinzu kommen Erkenntnisse zu den Motiven der Nachhaltigkeitsprüfung aus Sicht der Unternehmen. Sie deuten auf die Erhöhung der Glaubwürdigkeit und performance- sowie prozessbezogene Verbesserungen als Leitmotive hin (Edgley et al., 2010; Park & Brorson, 2005; Simnett et al., 2009), was eine positive Auslegung von Managerial Capture zumindest nicht ausschliesst.

Wird die externe Prüfung des Nachhaltigkeitsberichts als Teil der übergeordneten Nachhaltigkeitsstrategie eines Unternehmens betrachtet, so lässt sich annehmen, dass auch die Prüfung zur Sicherstellung der Legitimität der Unternehmensaktivitäten beitragen soll (Darnall et al., 2009; O'Dwyer et al., 2011). Damit kann die Prüfung letztlich neben der Nachhaltigkeitsberichterstattung selbst als weiteres strategisches Werkzeug zur Legitimitätssicherung gegenüber den Stakeholdern eines Unternehmens angesehen werden. Je nach Zielsetzung des Unternehmens, welches die Prüfung in Auftrag gibt und bezahlt, können sich daraus sowohl negative Effekte im Sinne des bisherigen Verständnisses von Managerial Capture als auch positive Effekte im Sinne einer ernsthaften Verbesserung interner Prozesse, der Nachhaltigkeitsperformance und der Transparenz gegenüber externen Stakeholdern ergeben.

⁸ Vgl. zum Beispiel die verpflichtende externe Prüfung in Frankreich seit 2012 (Grenelle II, 2010). Ebenso ist eine externe Prüfung in Schweden für Unternehmen im Staatsbesitz vorgeschrieben (Richtlinien zur externen Berichterstattung für staatliche Unternehmen, 2007). Werden einzelne Informationen mit Bezug zu Nachhaltigkeitsindikatoren im Geschäftsbericht veröffentlicht, so kann für diese ebenfalls eine Prüfungspflicht im Rahmen der Abschlussprüfung bestehen.

1.3 Beitrag und Zusammenfassung der eingebrachten Papers

Die drei in die kumulative Dissertation eingehenden Papers adressieren mehrere offene Fragen innerhalb der Forschungsgebiete I, II und III. Das erste Paper “Zusammenhang zwischen Nachhaltigkeitsperformance und Nachhaltigkeitsberichterstattung – Legitimität oder finanzielle Überlegungen?” kann dabei den Gebieten I und II zugeordnet werden. Diesem Paper liegt die Idee zu Grunde, dass die Nachhaltigkeitsberichterstattung von Unternehmen als strategisches Werkzeug eingesetzt wird, um ihren Marktwert zu Gunsten ihrer Shareholder zu steigern und gleichzeitig die Legitimität gegenüber anderen wesentlichen Stakeholdern sicherzustellen. Hinter beiden Zielen stehen mit der freiwilligen Offenlegungstheorie und der Legitimitätstheorie unterschiedliche theoretische Erklärungsansätze, die gemäss der bisherigen Literatur in Konkurrenz zueinander stehen und jeweils einen entgegengesetzten Zusammenhang zwischen Nachhaltigkeitsperformance und -berichterstattung implizieren. Nach der freiwilligen Offenlegungstheorie werden nur ausreichend positive Informationen berichtet, während nach der Legitimitätstheorie schlechte Informationen verschleiert und zur Ablenkung mit umfangreichen anderen Informationen ergänzt werden. Der erste Beitrag dieses Papers liegt in der neuen Sichtweise, dass beide Theorien gemeinsam und gleichzeitig Bestand haben können, da Unternehmen sowohl im Interesse ihrer Shareholder als auch zur Wahrung ihrer Legitimität im Interesse zentraler weiterer Stakeholder agieren können (Forschungsgebiet II). Eine Schlüsselrolle bei der Beseitigung des bisherigen Widerspruchs beider Theorien liegt in der Messung der Nachhaltigkeitsberichterstattung. Im Paper wird die Herangehensweise früherer Messkonzepte verfeinert, indem durchgängig zwischen einer quantitativen und somit vergleichbaren Form der Informationsbereitstellung und einer verbalen und somit schwer vergleichbaren Form unterschieden wird. So kann gezeigt werden, dass die marktorientierte freiwillige Offenlegungstheorie einen positiven Zusammenhang zwischen Nachhaltigkeitsperformance und quantitativer Berichterstattung erklären kann, während gleichzeitig die Legitimitätstheorie einen negativen Zusammenhang zwischen Performance und verbaler Berichterstattung erklärt. Bei der Messung der Berichterstattung und Performance kommen verbesserte Proxies zum Einsatz, die neben der ökologischen Dimension zusätzlich die soziale Dimension der Nachhaltigkeit berücksichtigen. Darin liegt der zweite Forschungsbeitrag dieses Papers (Forschungsgebiet I). Die empirischen Ergebnisse einer bivariaten Analyse für 50 Deutsche und Schweizer Grossunternehmen liefern erste Belege für die theoretisch vermuteten Zusammenhänge und zeigen dabei methodisch tiefer gehenden Forschungsbedarf an.

Das zweite Paper “The Relationship between Sustainability Performance and Sustainability Disclosure Quality – Hard Numbers Beat Smooth Talk” knüpft direkt am vom vorangegangenen Paper aufgezeigten Forschungsbedarf an. Damit die Vereinbarkeit von freiwilliger Offenlegungstheorie und Legitimitätstheorie noch klarer herausgearbeitet werden kann, blendet das Paper die Quantität der Nachhaltigkeitsberichterstattung aus und fokussiert speziell auf Unterschiede in der Qualität der berichteten Informationen. Die Stichprobe von 195 Unternehmen aus sieben voneinander abgegrenzten Europäischen Ländern und 19 individuell berücksichtigten Industriegruppen erlaubt belastbarere Aussagen zum Einfluss der Nachhaltigkeitsperformance auf die Nachhaltigkeitsberichterstattung. Dazu trägt auch der multivariate Analyseansatz mit weiteren unternehmensspezifischen Kontrollvariablen bei (Forschungsgebiet II). Das

Messkonzept zur Berichterstattung wird direkt auf deren Qualität zugeschnitten und berücksichtigt auch die industriespezifisch variierende Bedeutung einzelner Informationen. Ausserdem wird die Messung der Nachhaltigkeitsperformance im sozialen Bereich angepasst, um besser die tatsächliche Leistung der bewerteten Unternehmen abzubilden (Forschungsgebiet I). Die Ergebnisse stützen die Implikationen des ersten Forschungsbeitrags und zeigen, dass Unternehmen mit schlechter Nachhaltigkeitsperformance über die betroffenen Elemente aus Legitimitätsüberlegungen heraus in niedrigerer Qualität berichten, während vorwiegend gute Performer freiwillig Informationen von höherer Qualität offenlegen. Weiterhin lässt sich aus den Kontrollvariablen ablesen, dass eine in Fragen der unternehmerischen Nachhaltigkeit aktive strategische Orientierung, angezeigt durch die freiwillige externe Prüfung des Nachhaltigkeitsberichts, mit einer höheren Informationsqualität einhergeht. Bei gemeinsamer Betrachtung weisen das erste und das zweite Paper auf zentrale Probleme der Nachhaltigkeitsberichterstattung der letzten Jahre hin. Es wird deutlich, dass in mehreren Europäischen Ländern nicht nur eine Greenwashing-Problematik in Bezug auf schlechte Performer besteht, sondern dass auch hinsichtlich der Qualität der berichteten Informationen wesentliche Spielräume bestehen, innerhalb derer Unternehmensinteressen eine vergleichbare Nachhaltigkeitsberichterstattung verhindern können. Ohne mögliche unerwünschte Nebenwirkungen einer strengeren Regulierung der Nachhaltigkeitsberichterstattung zu betrachten, liefern die Papers zumindest theoretische Argumente zu Gunsten der aktuellen Entwicklung hin zu einheitlicheren und bindenden Vorschriften für Europäische Unternehmen.

Nachdem die beiden ersten Papers die Zusammenhänge zwischen Nachhaltigkeitsperformance und Berichterstattung sowie zwischen externer Prüfung und Berichterstattung aufgezeigt haben, beleuchtet das dritte Paper "The Nature of Managerial Capture in Sustainability Assurance Engagements" den noch fehlenden dritten Zusammenhang zwischen Nachhaltigkeitsperformance und der Prüfung von Nachhaltigkeitsberichten. Es geht um die Frage, wie das strategische Eigeninteresse der Unternehmen deren freiwillige und kostenpflichtige Auswahl externer Prüfungsleistungen beeinflusst. In der bisherigen Literatur wird davon ausgegangen, dass Managerial Capture zu einer beschönigenden und gegenüber externen Stakeholdern nicht ausreichend Transparenz schaffenden Prüfung führt. Das Paper stellt die Bedeutung der Prüfung als weiteres Werkzeug zur Legitimitätssicherung der Unternehmen heraus und leitet daraus ab, dass Managerial Capture nicht in jedem Fall zu einer eingeschränkten Prüfungsintensität führen muss. Die Ergebnisse der empirischen Untersuchung für 122 Europäische Unternehmen zeigen einen im Durchschnitt negativen Zusammenhang zwischen Nachhaltigkeitsperformance und Prüfungsintensität. Gerade die schlechten Performer profitieren am meisten von der zusätzlichen Glaubwürdigkeit ihrer Nachhaltigkeitsaktivitäten durch eine intensive externe Prüfung. Gleichzeitig kann die Prüfung zur Verbesserung interner Systeme und letztlich der Nachhaltigkeitsperformance beitragen. Somit kann Managerial Capture in einer regulierungsarmen Umgebung auch zu positiven Effekten führen, die nicht nur das geprüfte Unternehmen selbst betreffen, sondern durch den Zugewinn an Transparenz auch seine Stakeholder. Das Paper trägt neben seiner empirischen Perspektive und dem entwickelten Messkonzept der Prüfungsintensität mit einem breiteren und neutraleren Verständnis von Managerial Capture zum bisherigen Forschungsstand bei. Ausserdem werden mit der finanziellen Unternehmensperformance und den Gesamtaufwendungen für Leistungen von Prü-

fungsgesellschaften weitere Determinanten der Prüfungsintensität offengelegt, die auch im Verdrängungswettbewerb im Markt für die Prüfung von Nachhaltigkeitsberichten von Bedeutung sind. Mit dem positiven Einfluss sowohl der Unabhängigkeit als auch der Kompetenz des Prüfers wird zusätzlich die Anbieterseite in die Betrachtung einbezogen (Forschungsgebiet III).

Die übergreifende Betrachtung der Ergebnisse aller drei Papers macht deutlich, dass hinsichtlich der Nachhaltigkeitsberichterstattung und deren Prüfung bei Europäischen Unternehmen noch Entwicklungspotenzial besteht. Das heterogene Gesamtbild spiegelt nicht nur das theoretisch uneinheitliche Verständnis von Nachhaltigkeit im Unternehmenskontext wider, sondern zeigt auch die Vielfalt an Ausgestaltungsmöglichkeiten in einem vergleichsweise schwach regulierten Umfeld. Dem Thema umfassender Nachhaltigkeit wird von den Unternehmen weiterhin eine unterschiedliche strategische Bedeutung beigemessen, die von Minimalstandards bis hin zu einer Einstufung als zentralen langfristigen Erfolgsfaktor reicht. Es bleibt abzuwarten, ob sich in Zukunft mehr Unternehmen von den zumindest theoretisch implizierten Vorteilen eines umfassenden Nachhaltigkeitsverständnisses für die Unternehmen selbst und für ihre Stakeholder überzeugen lassen.

Literaturverzeichnis

- Aerts, W., & Cormier, D. (2009). Media legitimacy and corporate environmental communication. *Accounting, Organizations and Society*, 34(1), 1–27.
- Al-Tuwaijri, S. A., Christensen, T. E., & Hughes, K. E. (2004). The relations among environmental disclosure, environmental performance, and economic performance: A simultaneous equations approach. *Accounting, Organizations and Society*, 29(5–6), 447–471.
- Ameer, R., & Othman, R. (2012). Sustainability practices and corporate financial performance: A study based on the top global corporations. *Journal of Business Ethics*, 108(1), 61–79.
- Ball, A., Owen, D. L., & Gray, R. (2000). External transparency or internal capture? The role of third-party statements in adding value to corporate environmental reports. *Business Strategy and the Environment*, 9(1), 1–23.
- Bewley, K., & Li, Y. (2000). Disclosure of environmental information by Canadian manufacturing companies: A voluntary disclosure perspective. *Advances in Environmental Accounting & Management*, 1, 201–226.
- Branco, M. C., & Rodrigues, L. L. (2008). Factors influencing social responsibility disclosure by Portuguese companies. *Journal of Business Ethics*, 83(4), 685–701.
- Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? An empirical investigation of corporate environmental reputation. *Accounting, Organizations and Society*, 37(1), 14–25.

- Cho, C. H., Michelon, G., Patten, D. M., & Roberts, R. W. (2014). CSR report assurance in the USA: An empirical investigation of determinants and effects. *Sustainability Accounting, Management and Policy Journal*, 5(2), 130–148.
- Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, Organizations and Society*, 32(7–8), 639–647.
- Cho, C. H., Patten, D. M., & Roberts, R. W. (2006). Corporate political strategy: An examination of the relation between political expenditures, environmental performance, and environmental disclosure. *Journal of Business Ethics*, 67(2), 139–154.
- Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, 33(4–5), 303–327.
- Clarkson, P. M., Overell, M. B., & Chapple, L. (2011). Environmental reporting and its relation to corporate environmental performance. *Abacus*, 47(1), 27–60.
- Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, 3, 1–44.
- Cooper, S. M., & Owen, D. L. (2007). Corporate social reporting and stakeholder accountability: The missing link. *Accounting, Organizations and Society*, 32(7–8), 649–667.
- Cormier, D., & Magnan, M. (2003). Environmental reporting management: A continental European perspective. *Journal of Accounting and Public Policy*, 22(1), 43–62.
- Cormier, D., Magnan, M., & van Velthoven, B. (2005). Environmental disclosure quality in large German companies: Economic incentives, public pressures or institutional conditions? *European Accounting Review*, 14(1), 3–39.
- Darnall, N., Seol, I., & Sarkis, J. (2009). Perceived stakeholder influences and organizations' use of environmental audits. *Accounting, Organizations and Society*, 34(2), 170–187.
- Davis, K. (1973). The case for and against business assumption of social responsibilities. *Academy of Management Journal*, 16(2), 312–322.
- Dawkins, C., & Fraas, J. W. (2011). Coming clean: The impact of environmental performance and visibility on corporate climate change disclosure. *Journal of Business Ethics*, 100(2), 303–322.
- de Villiers, C., & van Staden, C. J. (2006). Can less environmental disclosure have a legitimizing effect? Evidence from Africa. *Accounting, Organizations and Society*, 31(8), 763–781.
- de Villiers, C., & van Staden, C. J. (2011). Where firms choose to disclose voluntary environmental information. *Journal of Accounting and Public Policy*, 30(6), 504–525.
- Deegan, C. (2002). Introduction: The legitimizing effect of social and environmental disclosures – A theoretical foundation. *Accounting, Auditing & Accountability Journal*, 15(3), 282–311.
- Deegan, C., Cooper, B. J., & Shelly, M. (2006). An investigation of TBL report assurance statements: UK and European evidence. *Managerial Auditing Journal*, 21(4), 329–371.

- Deloitte (2014). Anwendung der IFRS in einzelnen Rechtskreisen. http://www.iasplus.com/de/resources/ifrs-topics/use-of-ifrs?set_language=de. Letzter Zugriff am 03.12.2014.
- Deutsche Börse (2014). Börsenordnung für die Frankfurter Wertpapierbörse. http://www.deutsche-boerse.com/dbg/dispatch/de/binary/gdb_navigation/info_center?object_id=84XHGGZ360NSGDDE. Letzter Zugriff am 03.12.2014.
- Dowling, J., & Pfeffer, J. (1975). Organizational legitimacy: Social values and organizational behavior. *The Pacific Sociological Review*, 18(1), 122–136.
- Edgley, C. R., Jones, M. J., & Solomon, J. F. (2010). Stakeholder inclusivity in social and environmental report assurance. *Accounting, Auditing & Accountability Journal*, 23(4), 532–557.
- Elkington, J. (1997). *Cannibals with forks: The triple bottom line of 21st century business*. Oxford: Capstone.
- FASB (2014). International convergence of accounting standards – A brief history. <http://www.fasb.org/jsp/FASB/Page/SectionPage&cid=1176156304264>. Letzter Zugriff am 03.12.2014.
- Fekrat, M. A., Inclan, C., & Petroni, D. (1996). Corporate environmental disclosures: Competitive disclosure hypothesis using 1991 annual report data. *The International Journal of Accounting*, 31(2), 175–195.
- Fifka, M. (2012). The development and state of research on social and environmental reporting in global comparison. *Journal für Betriebswirtschaft*, 62(1), 45–84.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston, MA: Cambridge University Press.
- GRI (2011). Sustainability reporting guidelines version 3.1. <https://www.globalreporting.org/resourcelibrary/g3.1-guidelines-incl-technical-protocol.pdf>. Letzter Zugriff am 03.12.2014.
- GRI (2013a). G4 sustainability reporting guidelines: Reporting principles and standard disclosures. <https://www.globalreporting.org/resourcelibrary/GRIG4-Part1-Reporting-Principles-and-Standard-Disclosures.pdf>. Letzter Zugriff am 03.12.2014.
- GRI (2013b). G4 sustainability reporting guidelines: Implementation manual. <https://www.globalreporting.org/resourcelibrary/GRIG4-Part2-Implementation-Manual.pdf>. Letzter Zugriff am 03.12.2014.
- Hoffmann, T. (2011). *Unternehmerische Nachhaltigkeitsberichterstattung: Eine Analyse des GRI G3.1-Berichtsrahmens*. Lohmar: Eul.
- Hughes, S. B., Anderson, A., & Golden, S. (2001). Corporate environmental disclosures: Are they useful in determining environmental performance? *Journal of Accounting and Public Policy*, 20(3), 217–240.
- IASB (2014). IFRS application around the world: Jurisdiction profiles. <http://www.ifrs.org/Use-around-the-world/Pages/Jurisdiction-profiles.aspx>. Letzter Zugriff am 03.12.2014.

- Ihlen, Ø., & Roper, J. (2014). Corporate reports on sustainability and sustainable development: 'We have arrived'. *Sustainable Development*, 22(1), 42–51.
- Ingram, R. W., & Frazier, K. B. (1980). Environmental performance and corporate disclosure. *Journal of Accounting Research*, 18(2), 614–622.
- International Council on Mining and Metals (2014). Sustainable development framework. <http://www.icmm.com/our-work/sustainable-development-framework>. Letzter Zugriff am 03.12.2014.
- Kolk, A. (2003). Trends in sustainability reporting by the Fortune Global 250. *Business Strategy and the Environment*, 12(5), 279–291.
- Kolk, A., & Perego, P. (2010). Determinants of the adoption of sustainability assurance statements: An international investigation. *Business Strategy and the Environment*, 19(3), 182–198.
- Kolk, A., Walhain, S., & van de Wateringen, S. (2001). Environmental reporting by the Fortune Global 250: Exploring the influence of nationality and sector. *Business Strategy and the Environment*, 10(1), 15–28.
- Kollman, K., & Prakash, A. (2001). Green by choice? Cross-national variations in firms' responses to EMS-based environmental regimes. *World Politics*, 53(3), 399–430.
- KPMG (2013). The KPMG survey of corporate responsibility reporting 2013. <http://www.kpmg.com/Global/en/IssuesAndInsights/ArticlesPublications/corporate-responsibility/Documents/kpmg-survey-of-corporate-responsibility-reporting-2013.pdf>. Letzter Zugriff am 03.12.2014.
- Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics & Management Strategy*, 20(1), 3–41.
- Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good...and does it matter? A meta-analysis of the relationship between corporate social and financial performance. *Working paper: Harvard University et al.*
- Marquis, C., & Toffel, M. W. (2011). The globalization of corporate environmental disclosure: Accountability or greenwashing? *Working paper: Harvard University*.
- Marshall, J. D., & Toffel, M. W. (2005). Framing the elusive concept of sustainability: A sustainability hierarchy. *Environmental Science & Technology*, 39(3), 673–682.
- McGuire, J. W. (1969). The changing nature of business responsibilities. *The Journal of Risk and Insurance*, 36(1), 55–61.
- Neu, D., Warsame, H., & Pedwell, K. (1998). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society*, 23(3), 265–282.
- Norman, W., & MacDonald, C. (2004). Getting to the bottom of "triple bottom line". *Business Ethics Quarterly*, 14(2), 243–262.

- O'Donovan, G. (2002). Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. *Accounting, Auditing & Accountability Journal*, 15(3), 344–371.
- O'Dwyer, B. (2011). The case of sustainability assurance: Constructing a new assurance service. *Contemporary Accounting Research*, 28(4), 1230–1266.
- O'Dwyer, B., Owen, D., & Unerman, J. (2011). Seeking legitimacy for new assurance forms: The case of assurance on sustainability reporting. *Accounting, Organizations and Society*, 36(1), 31–52.
- O'Dwyer, B., & Owen, D. L. (2005). Assurance statement practice in environmental, social and sustainability reporting: A critical evaluation. *The British Accounting Review*, 37(2), 205–229.
- Orij, R. (2010). Corporate social disclosures in the context of national cultures and stakeholder theory. *Accounting, Auditing & Accountability Journal*, 23(7), 868–889.
- Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies*, 24(3), 403–441.
- Ott, K., & Döring, R. (2008). *Theorie und Praxis starker Nachhaltigkeit* (2. Aufl.). Marburg: Metropolis.
- Owen, D. L., Swift, T. A., Humphrey, C., & Bowerman, M. (2000). The new social audits: Accountability, managerial capture or the agenda of social champions? *The European Accounting Review*, 9(1), 81–98.
- Park, J., & Brorson, T. (2005). Experiences of and views on third-party assurance of corporate environmental and sustainability reports. *Journal of Cleaner Production*, 13(10–11), 1095–1106.
- Patten, D. M. (2002). The relation between environmental performance and environmental disclosure: A research note. *Accounting, Organizations and Society*, 27(8), 763–773.
- Pigou, A. C. (1932). *The economics of welfare* (4th ed.). London: Macmillan.
- PricewaterhouseCoopers (2014). IFRS adoption by country. http://www.pwc.com/en_US/us/issues/ifrs-reporting/publications/assets/pwc-ifrs-by-country-2014.pdf. Letzter Zugriff am 03.12.2014.
- Roberts, R. W. (1992). Determinants of corporate social responsibility disclosure: An application of stakeholder theory. *Accounting, Organizations and Society*, 17(6), 595–612.
- Rockness, J. W. (1985). An assessment of the relationship between US corporate environmental performance and disclosure. *Journal of Business Finance & Accounting*, 12(3), 339–354.
- Sethi, S. P. (1978). Advocacy advertising – The American experience. *California Management Review*, 21(1), 55–68.
- Simnett, R., Vanstraelen, A., & Chua, W. F. (2009). Assurance on sustainability reports: An international comparison. *The Accounting Review*, 84(3), 937–967.

- SIX Swiss Exchange (2014). Kotierungsreglement. http://www.six-exchange-regulation.com/admission_manual/03_01-LR_de.pdf. Letzter Zugriff am 03.12.2014.
- Suchman, M. C. (1995). Managing legitimacy: Strategies and institutional approaches. *Academy of Management Review*, 20(3), 571–610.
- Ullmann, A. A. (1985). Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of U.S. firms. *Academy of Management Review*, 10(3), 540–557.
- United Nations (1987). Report of the world commission on environment and development: Our common future. <http://www.un-documents.net/our-common-future.pdf>. Letzter Zugriff am 03.12.2014.
- van der Laan Smith, J., Adhikari, A., & Tondkar, R. H. (2005). Exploring differences in social disclosures internationally: A stakeholder perspective. *Journal of Accounting and Public Policy*, 24(2), 123–151.
- Verband der Chemischen Industrie, Industriegewerkschaft Bergbau, Chemie, Energie, & Bundesarbeitgeberverband Chemie (2013). Leitlinien zur Nachhaltigkeit für die chemische Industrie in Deutschland. https://www.chemiehoch3.de/fileadmin/user_upload/downloads/chemie-hoch-3-broschuere-leitlinien-de.pdf. Letzter Zugriff am 03.12.2014.
- Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, 5(1), 179–194.
- Wiseman, J. (1982). An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations and Society*, 7(1), 53–63.
- Wood, D. J. (1991). Corporate social performance revisited. *Academy of Management Review*, 16(4), 691–718.
- Wood, D. J. (2010). Measuring corporate social performance: A review. *International Journal of Management Reviews*, 12(1), 50–84.

Zusammenhang zwischen Nachhaltigkeitsperformance und Nachhaltigkeitsberichterstattung – Legitimität oder finanzielle Überlegungen?



Katrin Hummel und Christian Schlick

Nachhaltigkeitsberichterstattung, Nachhaltigkeitsperformance, Legitimacy Theory, Voluntary Disclosure Theory

Sustainability disclosure, sustainability performance, legitimacy theory, voluntary disclosure theory



Der Zusammenhang zwischen Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance ist zum aktuellen Zeitpunkt weder theoretisch noch empirisch eindeutig geklärt. Theoretische Erklärungsansätze liefern sowohl die Legitimacy Theory als auch die Voluntary Disclosure Theory. Angesichts methodischer Schwächen bisheriger Studien bei der Operationalisierung der zentralen Untersuchungsvariablen steht im Mittelpunkt des Beitrags die Entwicklung neuer Konzepte zur Messung von Umfang und Form der Nachhaltigkeitsberichterstattung sowie der Nachhaltigkeitsperformance. Die entwickelten Messkonzepte eignen sich zur konsistenten Überprüfung der theoretischen Erklärungsansätze. Anhand einer Stichprobe von 50 Deutschen und Schweizer Grossunternehmen kann im Unterschied zu bisherigen Untersuchungen gezeigt werden, dass beide Theorien jeweils einen Teil der Nachhaltigkeitsberichterstattung erklären können.

Both theoretical considerations as well as empirical findings report mixed results on the relationship between sustainability disclosure and sustainability performance. Theoretical considerations are mainly based on legitimacy theory and voluntary disclosure theory. Due to shortcomings of previous studies with regard to the measurement of the main variables, our paper concentrates on new measurement approaches for sustainability disclosure and sustainability performance. Findings from a sample of 50 German and Swiss companies support our methodological approach. In particular, we find empirical evidence that legitimacy theory and voluntary disclosure theory likewise can explain certain aspects of sustainability reporting.

1. Problemstellung

Die Frage, warum Unternehmen freiwillig Informationen zur Nachhaltigkeitsperformance im Rahmen eines Geschäftsberichts oder eines separaten Nachhaltigkeitsberichts veröffentlichen, ist zum aktuellen Zeitpunkt weder theoretisch noch empirisch eindeutig beant-

wortet. Von besonderem Interesse ist dabei der Zusammenhang zwischen Nachhaltigkeitsperformance und Nachhaltigkeitsberichterstattung, wobei zwei (vermeintlich) divergierende theoretische Erklärungsansätze diskutiert werden. Der *Legitimacy Theory* (LT) folgend wird argumentiert, dass die Nachhaltigkeitsberichterstattung eingesetzt wird, um die öffentliche Wahrnehmung der (tatsächlich schlechten) Nachhaltigkeitsperformance eines Unternehmens positiv zu beeinflussen (vgl. exemplarisch Deegan 2002). Daher wird in empirischen Untersuchungen ein *negativer* Zusammenhang zwischen der Nachhaltigkeitsperformance und dem Umfang der Nachhaltigkeitsberichterstattung oftmals mit Argumenten der LT erklärt (vgl. exemplarisch Dawkins/Fraas 2011, 315f. und Patten 2002, 772). Demgegenüber argumentieren Vertreter der *Voluntary Disclosure Theory* (VDT) mit preissteigernden Effekten positiver Nachrichten am Kapitalmarkt und postulieren einen *positiven* Zusammenhang zwischen der Nachhaltigkeitsperformance und dem Umfang bzw. der Qualität der Nachhaltigkeitsberichterstattung (vgl. exemplarisch Clarkson et al. 2008).

Die Ergebnisse bisheriger empirischer Untersuchungen mehrheitlich US-amerikanischer Unternehmen fallen konträr aus und sind insofern wenig geeignet, die Fragestellung eindeutig zu beantworten. Sie konzentrieren sich ausschliesslich auf den Umweltbereich als einen Teilbereich der Nachhaltigkeit. Insbesondere weisen sämtliche Untersuchungen bei der Messung und Operationalisierung der zentralen Variablen Schwächen auf, welche die Interpretation der Ergebnisse stark einschränken. So erfolgt nur bei wenigen Untersuchungen eine Unterscheidung zwischen quantitativer und verbaler Form der Berichterstattung. Diese Trennung ist jedoch zur Analyse der Problemstellung notwendig. Denn der Logik der LT folgend wird bei schlechter Nachhaltigkeitsperformance zwar umfangreicher berichtet, die Darstellung erfolgt aber bei kritischen Themen eher verbal. Hingegen basiert die VDT auf der Entscheidungsrelevanz der publizierten Informationen, was ebenfalls eine Unterscheidung von verbaler und quantitativer Information nahelegt. Des Weiteren fällt eine grosse Heterogenität bei der Messung und Operationalisierung der Nachhaltigkeitsperformance auf. Zur Operationalisierung wird mehrheitlich auf lediglich ein oder zwei isolierte Kenngrössen, wie beispielsweise Emissionsdaten oder Daten zur Abfallmenge, zurückgegriffen (vgl. exemplarisch Clarkson et al. 2008). Eine solche Messung der Umweltperformance ist hinsichtlich ihrer Validität fragwürdig und daher für eine grundlegende Beurteilung der Umweltperformance wenig geeignet.

Vor diesem Hintergrund besteht das Ziel dieses Beitrags in der Entwicklung und Anwendung von Konzepten zur Messung von Umfang und Form der Nachhaltigkeitsberichterstattung sowie der Nachhaltigkeitsperformance. Im Unterschied zu bisherigen Untersuchungen wird bei der Messung der Nachhaltigkeitsberichterstattung streng zwischen quantitativer und verbaler Form der Berichterstattung differenziert. Es wird ein valides Konzept zur Messung der Nachhaltigkeitsperformance vorgestellt. Dieses bietet wertvolle Ansatzpunkte zur Überwindung der Schwächen bisheriger Ansätze. Die Anwendung dieser Messkonzepte erlaubt belastbare Rückschlüsse auf die Gültigkeit der theoretischen Erklärungsansätze. Für ein Sample von 50 Deutschen und Schweizer Grossunternehmen wird gezeigt, dass beide Theorien jeweils einen *Teil* der Nachhaltigkeitsberichterstattung erklären können.

Der Beitrag gliedert sich in fünf Kapitel. Aufbauend auf einem umfassenden Überblick über bisherige empirische Untersuchungen des Zusammenhangs von Nachhaltigkeitsberichterstattung und -performance und einer Darstellung der theoretischen Erklärungsansätze

ze werden im zweiten Kapitel die Forschungshypothesen der Untersuchung formuliert. Diese bilden den Ausgangspunkt für die Entwicklung von Messkonzepten zur Bestimmung von Umfang und Form der Nachhaltigkeitsberichterstattung sowie der Nachhaltigkeitsperformance, die im dritten Kapitel vorgestellt und im Hinblick auf die Erfüllung zentraler Gütekriterien kritisch reflektiert werden. Das vierte Kapitel beinhaltet die empirische Anwendung der Messkonzepte sowie die Überprüfung der formulierten Forschungshypothesen. Der Beitrag schliesst mit einem Fazit und einer kritischen Diskussion.

2. Aktueller Forschungsstand, theoretische Fundierung und Hypothesenformulierung

Tabelle 1 gibt einen Überblick über bisherige empirische Untersuchungen des Zusammenhangs von Nachhaltigkeitsberichterstattung und -performance. Gemäss des Triple-Bottom-Line-Ansatzes umfasst Nachhaltigkeit die drei Säulen Ökonomie, Ökologie und Soziales (vgl. *Elkington* 1999). Im Kern geht es um die ökonomische, ökologische und soziale Verantwortung, welche eine Unternehmung gegenüber ihrem Umfeld wahrzunehmen hat.

Autor (Jahr)	n Land	Messung Berichterstattung	Messung Performance	Ergebnis
<i>Cho et al.</i> (2012)	92 US	<ul style="list-style-type: none"> Analyse Nachhaltigkeits- und Geschäftsberichte vgl. <i>Clarkson et al.</i> (2008) 	<ul style="list-style-type: none"> Environmental Impact Score von Newsweek: u.a. Luftemissionen, Wassernutzung, Abfälle 	-
<i>Clarkson et al.</i> (2011)	51 AU	<ul style="list-style-type: none"> Analyse Nachhaltigkeits- und Geschäftsberichte vgl. <i>Clarkson et al.</i> (2008) 	<ul style="list-style-type: none"> giftige Freisetzen pro Umsatz (90 risikogewichtete Substanzen) 	-
<i>Dawkins/Fraas</i> (2011)	344 US	<ul style="list-style-type: none"> Reaktionen auf Umfrage des Carbon Disclosure Projects 3 Gruppen: keine Antwort/Ablehnung, unvollständige Antwort/Verweise, vollständige Beantwortung 	<ul style="list-style-type: none"> KLD-Rating: Stärken und Schwächen Rating der CO₂-Emissionen 	-
<i>Clarkson et al.</i> (2008)	191 US	<ul style="list-style-type: none"> Analyse Nachhaltigkeitsberichte 7 Kategorien mit 95 Elementen in Anlehnung an GRI (16 soft/leicht imitierbar und 79 hard/schwer imitierbar) 0/1-Punkteschema 	<ul style="list-style-type: none"> Recyclingrate giftiger Abfälle (Rang) Giftige Abfälle pro Umsatz (Rang) 	+
<i>Cho/Patten</i> (2007)	100 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte vgl. <i>Patten</i> (2002) zusätzlich: getrennte Auswertung der 4 finanziellen und 4 nicht-finanziellen Elemente 	<ul style="list-style-type: none"> KLD-Rating: Einteilung in schlechter/besser, sofern Schwächen im Umweltbereich identifiziert/nicht identifiziert werden 	-
<i>Cho et al.</i> (2006)	119 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte vgl. <i>Patten</i> (2002) 	<ul style="list-style-type: none"> KLD-Rating: Punktschwere über umweltbezogene Schwächen 	-
<i>de Villiers/van Staden</i> (2006)	≥137 ZA	<ul style="list-style-type: none"> Analyse Geschäftsberichte in Anlehnung an <i>Wiseman</i> (1982) Trennung allgemeine/spezifische Angaben 0/1-Punkteschema ohne Gewichtung getrennte Betrachtung von 8 allgemeinen Elementen der mitarbeiterbezogenen Berichterstattung 	<ul style="list-style-type: none"> keine explizite Messung Zusammenhang wird unterstellt 	-

Autor (Jahr)	n Land	Messung Berichterstattung	Messung Performance	Ergebnis
<i>Al-Tuwaijri et al.</i> (2004)	198 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte 4 Elemente: giftiger Abfall, Geldstrafen, Beseitigung giftiger Altlasten, Öl- und Chemieunfälle Punkteschema: quantitativ (3), verbal spezifisch (2), verbal unspezifisch (1), keine Berichterstattung (0) Relevanz für individuelles Unternehmen wird berücksichtigt 	<ul style="list-style-type: none"> Recyclingrate giftiger Abfälle Giftiger Abfall pro Umsatz Haltung zu Umweltbelangen (3 Faktoren) 	+
<i>Patten</i> (2002)	131 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte 4 verbale (Diskussion) und 4 quantitative (Kosten) Elemente 0/1-Punkteschema zusätzlich: Anzahl publizierter Zeilen 	<ul style="list-style-type: none"> Menge freigesetzter gefährlicher Stoffe pro Umsatz 	-
<i>Hughes et al.</i> (2001)	51 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte vgl. <i>Wiseman</i> (1982) zusätzlich 5 neue Elemente Punkteschema: quantitativ (4), verbal (3), oberflächlich (2), unerheblich (1), keine (0) 	<ul style="list-style-type: none"> 3 Gruppen (gut, gemischt, schlecht) auf Basis eines Rankings des CEP (Council on Economic Priorities): Emissionen in Luft und Wasser, Vermeidungsprogramme 	-
<i>Bewley/Li</i> (2000)	188 CA	<ul style="list-style-type: none"> Analyse Geschäftsberichte vgl. <i>Wiseman</i> (1982) zusätzlich: Trennung finanzielle/nicht-finanzielle Angaben 	<ul style="list-style-type: none"> Verschmutzungsneigung (hoch/niedrig je nach Branche) Meldung giftiger Freisetzen (ja/nein) 	+
<i>Fekrat et al.</i> (1996)	26 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte vgl. <i>Wiseman</i> (1982) 	<ul style="list-style-type: none"> vgl. <i>Ingram/Frazier</i> (1980) 	?
<i>Rockness</i> (1985)	26 US	<ul style="list-style-type: none"> indirekte Messung als Wahrnehmung der Angaben in Geschäftsberichten durch Probanden 	<ul style="list-style-type: none"> vgl. <i>Ingram/Frazier</i> (1980) 	?
<i>Wiseman</i> (1982)	26 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte 4 Kategorien mit 18 Elementen: ökonomisch, Rechtsstreitigkeiten, Verschmutzung, andere Angaben Punktbewertungsschema: quantitativ (3), verbal (2), allgemein (1), keine Berichterstattung (0) zusätzlich: Anzahl publizierter Zeilen 	<ul style="list-style-type: none"> vgl. <i>Ingram/Frazier</i> (1980) 	?
<i>Ingram/Frazier</i> (1980)	40 US	<ul style="list-style-type: none"> Analyse Geschäftsberichte 4 Kategorien mit insgesamt 20 Ausprägungen: Art (u.a. finanziell, nicht-finanziell, qualitativ), Zeitbezug (Vergangenheit, Gegenwart, Zukunft), Genauigkeit (spezifisch, allgemein), Thema Einzelprüfung Satz für Satz Auswertung je Ausprägung und je Kategorie (Summe) 	<ul style="list-style-type: none"> Index des CEP (Council on Economic Priorities): Emissionen in Luft und Wasser, Vermeidungsprogramme 	?

AU Australien CA Kanada US USA ZA Südafrika

+ positiver Zusammenhang - negativer Zusammenhang ? unklarer Zusammenhang

Tabelle 1: Überblick über bisherige empirische Untersuchungen des Zusammenhangs von Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance

Sämtliche Untersuchungen in *Tabelle 1* konzentrieren sich auf den Umweltbereich und nehmen keine umfassende Analyse der Nachhaltigkeitsperformance vor. Zwar gibt es auch empirische Untersuchungen der Nachhaltigkeitsberichterstattung, allerdings bleibt bei diesen Untersuchungen die Nachhaltigkeitsperformance unberücksichtigt (vgl. exemplarisch *Quick/Knocinski 2006; van der Laan Smith et al. 2005*). Für die Ableitung möglicher Zusammenhänge zwischen Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance werden in der Regel die Erklärungsansätze der VDT oder der LT herangezogen.

Gemäss der VDT, welche sich ursprünglich auf die Finanzberichterstattung am Kapitalmarkt bezieht, wird das freiwillige Publikationsverhalten von Unternehmen durch die Information selbst und die anfallenden Kosten der Publizität bestimmt (vgl. *Verrecchia 1983*). Die publizierte Information schlägt sich am Markt direkt im Unternehmenswert nieder. Es kann formal eine kritische Schwelle bestimmt werden, bei der die anfallenden Kosten der Publizität exakt der erwarteten Änderung des Unternehmenswertes aufgrund der Informationsbekanntgabe entsprechen. Informationen oberhalb dieser kritischen Schwelle werden berichtet, Informationen unterhalb dieser kritischen Schwelle werden nicht ausgewiesen. Zusammengefasst bietet das Modell eine Erklärung dafür, warum es aus Sicht eines Unternehmens rational sein kann, Informationen ab einem bestimmten Grenzwert ihrer Ausprägung nicht zu berichten. Diese Argumentationslogik wurde von verschiedenen Vertretern auf die freiwillige Berichterstattung nicht-finanzieller Informationen erweitert (vgl. exemplarisch *Clarkson et al. 2008; Bewley/Li 2000; Li et al. 1997*). Es wird argumentiert, dass nicht-finanzielle Informationen immer dann berichtet werden, wenn deren preissteigernde Effekte grösser sind als die damit verbundenen Kosten der Berichterstattung. Im Mittelpunkt steht somit die Entscheidung der Berichterstattung an sich, nicht jedoch in welcher Form berichtet wird. Der Argumentation der VDT folgend wird daher *Hypothese H₁* formuliert:

H₁: Es besteht ein positiver Zusammenhang zwischen der Nachhaltigkeitsperformance und dem Umfang berichteter Nachhaltigkeitsinformationen.

Implizit wird bei der VDT Entscheidungsrelevanz der berichteten Informationen unterstellt. In Anlehnung an Studien jüngerer Datums (vgl. *Clarkson et al. 2008; Cho/Patten 2007; Al-Tuwaijri et al. 2004; Bewley/Li 2000*) wird daher in einem zweiten Schritt die berichtete Nachhaltigkeitsinformation dahingehend differenziert, ob es sich um eine verbale oder quantitative Angabe handelt. Es wird argumentiert, dass insbesondere quantitative Informationen entscheidungsrelevant sind, da diese im Vergleich zu einer rein verbalen Berichterstattung eine bessere Objektivität und Vergleichbarkeit sowie eine geringere Manipulationsanfälligkeit aufweisen (vgl. *Al-Tuwaijri et al. 2004, 454*). Um adverse Selektion zu vermeiden, signalisieren Unternehmen mit guter Nachhaltigkeitsperformance ihren Zustand, indem sie quantitative Nachhaltigkeitsinformationen publizieren (vgl. *Clarkson et al. 2008, 307f.*). So können sie sich von Unternehmen mit schlechter Nachhaltigkeitsperformance unterscheiden. Zudem kann die Publikation einer quantitativen Information von Unternehmen mit schlechter Nachhaltigkeitsperformance nicht ohne Weiteres imitiert werden (vgl. *Clarkson et al. 2008, 308*). Diesem Erklärungsansatz folgend wird daher vermutet, dass Unternehmen mit besserer Nachhaltigkeitsperformance Informationen vermehrt quantitativ berichten. Diese Argumentation schlägt sich in folgender Hypothese nieder:

H_{2a}: Es besteht ein positiver Zusammenhang zwischen der Nachhaltigkeitsperformance und dem Ausmass der quantitativen Nachhaltigkeitsberichterstattung.

Einen anderen Erklärungsansatz bietet die LT, wonach der Fortbestand eines Unternehmens von der Legitimierung durch sein Umfeld abhängig ist (vgl. Davis 1973, 314). Dabei beschreibt Legitimität einen Zustand, bei welchem sich die Wertevorstellungen von Unternehmen und gesellschaftlichem Umfeld decken (vgl. Dowling/Pfeffer 1975, 122). Umgekehrt verwendet Sethi (1978, 58) den Begriff der Legitimitätslücke, sofern diese voneinander abweichen und bestimmte Unternehmenstätigkeiten von der Gesellschaft als Normverletzung angesehen werden. Eine Wahrnehmung als Normverletzung ist jedoch auch dann möglich, wenn die Wertvorstellungen von Unternehmen und Gesellschaft übereinstimmen. Durch Informationsasymmetrien zwischen Unternehmen und Gesellschaft könnte die Gesellschaft eine tatsächlich wertkonforme Unternehmenstätigkeit fälschlicherweise als Normverletzung ansehen. In diesem Zustand der Legitimitätslücke kann dem Unternehmen die Versorgung mit personellen und materiellen Ressourcen durch die Gesellschaft erschwert oder verweigert werden. Um eine Legitimitätslücke zu schliessen, soll das Unternehmen durch die gezielte Bereitstellung und vorteilhafte Darstellung neuer Informationen die vom Umfeld wahrgenommene Unternehmensleistung zu Gunsten des Unternehmens beeinflussen (vgl. Milne/Patten 2002, 381; Gray et al. 1995, 54; Sethi 1978, 58). Nur wenn diese Strategie nicht erfolgreich ist, wird die Anpassung der tatsächlichen Unternehmenstätigkeit empfohlen (vgl. Sethi 1978, 58). Neben der Wiederherstellung (repair) von Legitimität existieren weitere Überlegungen zu ihrem Aufbau (gain) und Erhalt (maintain) (vgl. O'Donovan 2002, 349; Suchman 1995, 585ff.), wobei einem Unternehmen der präventive Erhalt der Legitimität nahegelegt wird (vgl. Ashforth/Gibbs 1990, 183). Übertragen auf die Nachhaltigkeitsberichterstattung bedeutet dies, dass Unternehmen Informationen zur Nachhaltigkeit ihrer Tätigkeit als ein Mittel zur gezielten Beeinflussung ihrer Legitimität nutzen können (vgl. Deegan 2002). Es wird vermutet, dass eine solch gezielte Beeinflussung der von der Gesellschaft wahrgenommenen Nachhaltigkeitsperformance insbesondere durch Publikation *verbaler* Informationen erfolgt, da nur bei diesen eine gezielt vorteilhafte Darstellung überhaupt möglich ist. Sowohl die Vergleichbarkeit als auch die Objektivität dieser Form der Information sind stark eingeschränkt. Je schlechter die Nachhaltigkeitsperformance eines Unternehmens ausfällt, desto eher wird es die klar nachvollziehbare quantitative Berichterstattung bei kritischen Sachverhalten durch eine verbale Darstellung ersetzen, um eine existierende oder drohende Legitimitätslücke zu schliessen. Dies führt zu folgender Hypothese:

H_{2b}: Es besteht ein negativer Zusammenhang zwischen der Nachhaltigkeitsperformance und dem Ausmass der verbalen Nachhaltigkeitsberichterstattung.

Eine Analyse der Studien aus *Tabelle 1* zeigt, dass in bisherigen Studien die Messung und Operationalisierung der dort interessierenden Variablen Umweltberichterstattung und Umweltperformance sehr unterschiedlich vorgenommen werden.

Für die Messung der *Umweltberichterstattung* hat sich die Inhaltsanalyse als methodischer Standard herausgebildet, wenngleich die berücksichtigten Indikatoren und die Art der Erfassung stark variieren. Für die Auswertung wird in der Regel auf eine aggregierte Gesamtbewertung zurückgegriffen. Lediglich bei einigen Studien jüngerer Datums werden ansatzweise quantitative und qualitative Information unterschieden (vgl. Clarkson et al. 2008; Cho/Patten 2007; Al-Tuwarijri et al. 2004; Bewley/Li 2000). Allerdings ist diese

Unterscheidung weder bei der Messung noch bei der Auswertung durchgängig umgesetzt.¹ Dies könnte darin begründet sein, dass bisherige empirische Untersuchungen darauf abzielen, die Gültigkeit entweder der einen *oder* der anderen theoretischen Argumentation zu festigen, indem der Zusammenhang zwischen Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance untersucht wird. Ein positiver Zusammenhang wird als Beleg für die Gültigkeit der VDT, ein negativer Zusammenhang als Beleg für die Gültigkeit der LT interpretiert. Bei dieser Herangehensweise bleibt unbeachtet, dass beide Theorien für die Erklärung unterschiedlicher Aspekte der Nachhaltigkeitsberichterstattung gleichermaßen Bestand haben können. So schliessen *Clarkson et al.* (2008, 325) aus den Ergebnissen ihrer empirischen Untersuchung, dass die LT zusätzlich zur VDT geeignet zu sein scheint, die Umweltberichterstattung in bestimmten Situationen zu erklären.

Deutlich heterogener fällt die Messung der *Umweltperformance* aus. Häufig wird zur Operationalisierung auf lediglich ein oder zwei isolierte Kenngrößen, wie beispielsweise Emissionsdaten oder Daten zur Abfallmenge, zurückgegriffen. Ein solcher Ansatz erscheint im Hinblick auf die Validität der Messung fragwürdig. Alternativ wird in einigen Studien zur Messung der Umweltperformance auf die Beurteilung durch externe Ratingagenturen zurückgegriffen, wobei das Rating von *Kinder, Lydenberg, Domini* (KLD) am weitesten verbreitet ist (vgl. *Dawkins/Fraas* 2011; *Cho/Patten* 2007; *Cho et al.* 2006). Beim KLD-Rating wird die Umweltperformance eines Unternehmens anhand von 14 Indikatoren beurteilt. Pro Unternehmen werden sieben umweltbezogene Stärken und umweltbezogene Schwächen analysiert und binär (zutreffend/nicht zutreffend) codiert (vgl. *Chatterji et al.* 2009, 132ff.). Die Aggregation dieser Indikatoren ergibt ein Mass für die Umweltperformance eines Unternehmens. Auch wenn das KLD-Rating die Problematik der isolierten Berücksichtigung einzelner Aspekte der Umweltperformance überwindet, offenbart es gleichwohl Schwächen. Diese liegen zum einen in der Intransparenz, da konkrete, quantitative Messwerte sowie die Grenzwerte der binären Bewertung nicht nachvollziehbar sind. Zum anderen stellt die binäre Erfassung an sich eine starke Vereinfachung dar, da Streuungen innerhalb der Gruppen nicht erfasst werden. Ausserdem liegen die Ratings von KLD überwiegend für US-amerikanische Unternehmen und weniger für europäische Unternehmen vor.

Die knappe Diskussion der methodischen Ansätze bisheriger Untersuchungen zeigt, dass sich in der Forschung bislang keine Standards für eine differenzierte Messung von Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance herausgebildet haben. Bisherige Ansätze konzentrieren sich überwiegend auf den Umweltbereich und weisen zudem Schwächen auf, weshalb sie für die Hypothesenprüfung dieser Untersuchung ungeeignet sind. Im folgenden Abschnitt werden daher erweiterte Messkonzepte für die Nachhaltigkeitsberichterstattung und die Nachhaltigkeitsperformance entwickelt.

3. Untersuchungsmethodik

3.1 Gütekriterien der Messung

Die Messung von Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance sollte sich an den in der einschlägigen Literatur angeführten Gütekriterien Objektivität, Reliabi-

¹ Eine Ausnahme hiervon stellt die Untersuchung von *Cho und Patten* (2007) dar, in welcher der Zusammenhang zwischen Umweltperformance und Umweltberichterstattung für monetäre und nicht-monetäre Berichterstattung getrennt untersucht wird.

lität und Validität eines Messinstruments orientieren (vgl. Bortz/Döring 2006, 195ff.). Darüber hinaus erfordert der spezifische Untersuchungskontext eine Kongruenz der den beiden Messkonzepten zugrunde liegenden Informationsitems sowie eine mehrdimensionale Messung. Das Kriterium der *Objektivität* verlangt nach der Unabhängigkeit der Messung der interessierenden Variablen vom Beobachter (vgl. für die folgenden Ausführungen Bortz/Döring 2006, 195ff.). Mithin sollen verschiedene Anwender zu den gleichen Resultaten gelangen. Die *Reliabilität* ist ein Mass für die Zuverlässigkeit und Stabilität eines Messinstruments. Konkret sollen wiederholte Messungen dieselben Werte liefern. Die *Validität* beschreibt die Gültigkeit eines Messinstruments und untersucht, ob ein Messinstrument tatsächlich das misst, was gemessen werden soll. Die Anforderung der *inhaltlichen Kongruenz* bezieht sich auf die Gegenüberstellung von Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance. Insbesondere im Hinblick auf die Überprüfung der VDT ist nach Möglichkeit eine Deckungsgleichheit der erhobenen Informationsitems der beiden Messkonzepte erforderlich. Das Kriterium der *Mehrdimensionalität* folgt aus dem Triple-Bottom-Line-Ansatz und verlangt, dass möglichst alle Dimensionen der Nachhaltigkeit sowie verschiedene Aspekte der einzelnen Dimensionen der Nachhaltigkeit Berücksichtigung finden. In der vorliegenden Untersuchung erfolgt eine Konzentration auf die ökologische und soziale Dimension der Nachhaltigkeit. Dies ist im Wesentlichen zwei Überlegungen geschuldet. Zum einen liegen bereits zahlreiche Untersuchungen im Bereich der Finanzberichterstattung vor (für einen Überblick vgl. Healy/Palepu 2001, 405ff.). Zum anderen existieren gesetzliche Regelungen zur Bewertung und Berichterstattung ökonomischer Sachverhalte, so dass der Einbezug dieser Dimension die Ergebnisse einer Analyse des Zusammenhangs zwischen der freiwilligen Nachhaltigkeitsberichterstattung und -performance verzerren könnte.

3.2 Messung von Umfang und Form der Nachhaltigkeitsberichterstattung

In der vorliegenden Untersuchung wird *Nachhaltigkeitsberichterstattung* als Gesamtheit aller von einem Unternehmen der Öffentlichkeit absichtlich zugänglich gemachten Informationen über die ökonomischen, ökologischen und sozialen Aspekte seiner Entscheidungen abgegrenzt. Aus den in Abschnitt 3.1 genannten Gründen werden in dieser Arbeit die ökologischen und sozialen Aspekte herausgegriffen. In bisherigen empirischen Studien, die in *Tabelle 1* dargestellt sind, wird die Messung der Umweltberichterstattung überwiegend mittels Inhaltsanalyse vorgenommen. Hierbei wird die Berichterstattung zu definierten Informationsitems binär oder ordinal skaliert erfasst und die Einzelwerte werden zu einem Gesamtwert aggregiert (vgl. exemplarisch Clarkson *et al.* 2008 für eine binär skalierte und Al-Tuwaijri *et al.* 2004 für eine ordinal skalierte Messung). Ohne eine Unterscheidung der publizierten Informationen in quantitativ und verbal ist ein solches Messkonzept für die Beantwortung der Fragestellung nicht geeignet. Bei Binärcodierung spiegelt die Bewertung lediglich den Umfang der Berichterstattung wider und erlaubt demzufolge keine Aussagen zur Gültigkeit der Theorien. Selbst wenn eine präzise Unterscheidung in quantitativ/verbal auf Basis der Informationsitems vorgenommen wird, besteht die Gefahr, dass sich diese bei der Gesamtbewertung gegenseitig kompensieren. Diese Probleme können zwar durch eine in verbale Berichterstattung (gering) und quantitative Berichterstattung (hoch) ordinal trennende Bewertung überwunden werden, allerdings sind auch hier die Ergebnisse nicht eindeutig interpretierbar. So könnte ein positiver Zusammenhang sowohl als Beleg für die Gültigkeit der VDT herangezogen werden, wonach eine bessere Umweltperformance mit

einer vermehrt quantitativen Berichterstattung einhergeht, als auch als Beleg für die Gültigkeit der LT, wonach eine schlechtere Umweltperformance mit einer geringeren quantitativen und vermehrt verbalen Berichterstattung einhergeht.

Lediglich bei einigen Studien jüngerer Datums wird eine Unterscheidung von quantitativer und verbaler Form der Berichterstattung ansatzweise vorgenommen (vgl. *Clarkson et al.* 2008; *Cho/Patten* 2007; *Al-Tuwarijri et al.* 2004; *Bewley/Li* 2000). Ausgangspunkt für eine differenzierte Messung der Berichterstattung bei diesen Studien stellt die Zuordnung von Informationsitems zu einer harten/monetären oder einer weichen/nicht-monetären Berichterstattung dar. Zur Kategorie der weichen bzw. nicht-monetären Informationsitems werden in der Regel Angaben zu Strategie, Umweltinitiativen u.a. gezählt, wohingegen harte bzw. monetäre Informationsitems zentrale Umweltindikatoren wie Energieverbrauch oder Treibhausgasemissionen umfassen. Bei einem solchen Messansatz ist allerdings das Kriterium der inhaltlichen Kongruenz der Informationsitems nicht erfüllt. Sowohl bei Binärcodierung als auch bei einer ordinal skalierten Bewertung sind die Ergebnisse für die harte/monetäre Kategorie nicht eindeutig interpretierbar, da ein positiver Zusammenhang sowohl als Beleg für die Gültigkeit der LT (je schlechter die Umweltperformance, desto weniger harte Informationen werden berichtet) als auch als Beleg für die Gültigkeit der VDT (je besser die Umweltperformance, desto mehr harte Informationen werden berichtet) gedeutet werden kann. Dementsprechend fordern *Clarkson et al.* (2008, 325), „that future environmental disclosure research should move the focus of enquiry beyond the level of disclosure.“

Vor diesem Hintergrund soll im Folgenden ein Konzept zur Messung der Nachhaltigkeitsberichterstattung vorgestellt werden, welches für die Untersuchung der Problemstellung geeignet ist. Die in einem ersten Schritt erforderliche Auswahl der relevanten Informationselemente orientiert sich an den Leistungsindikatoren der aktuell gültigen Richtlinien der Global Reporting Initiative (GRI)² G3.1, welche weltweit quasi einen Standard der Nachhaltigkeitsberichterstattung darstellen (vgl. *Hoffmann* 2011, 68).³ Es werden jeweils zehn Einzelindikatoren für die Kategorien Ökologie und Soziales erfasst, wobei drei Leistungsindikatoren der Kategorie Ökologie für Unternehmen bestimmter Branchen durch branchenspezifische Angaben ersetzt werden. Zudem erfolgt je nach Bedeutung der Einzelindikatoren eine Gewichtung mit dem Faktor 1 oder 2. *Tabelle 2* zeigt die berücksichtigten Leistungsindikatoren in den einzelnen Kategorien mit dem jeweiligen Bezug zu den Leistungsindikatoren der GRI-Richtlinien und dem Gewichtungsfaktor. Die mit * gekennzeichneten Informationselemente werden bei Unternehmen der entsprechenden Branche durch die angeführten branchenspezifischen Indikatoren ersetzt.

2 Bei der GRI handelt es sich um eine Nichtregierungsorganisation, welche im Jahre 1997 mit dem Ziel gegründet wurde, die weltweite Verbreitung und Standardisierung der Nachhaltigkeitsberichterstattung voranzutreiben. Vgl. *Global Reporting Initiative* (2012).

3 Diese Vorgehensweise wird auch in anderen Studien gewählt, vgl. beispielsweise *Cho et al.* (2012), *Clarkson et al.* (2011) oder *Clarkson et al.* (2008).

Abkürzung	Leistungsindikator	Bezug zu G3.1	Faktor
Ökologie			
R-B1	Eingesetztes Material nach Gewicht oder Volumen	EN1	2
R-B2	Anteil Recyclingmaterial am Gesamtmaterialeinsatz	EN2	1
R-B3	Direkter und indirekter Energieverbrauch	EN3/4	2
R-B4	Gesamtwasserverbrauch nach Art	EN8	2
R-B5	Gesamte direkte und indirekte Treibhausgasemissionen	EN16	2
R-B6*	Emissionen von Ozon abbauenden Stoffen	EN19	1
R-B7	Andere Luftemissionen (Stickoxide, Schwefeloxide u.a.)	EN20	1
R-B8*	Gesamte Abwassereinleitungen nach Art und Einleitungsort	EN21	2
R-B9	Gesamtgewicht des Abfalls nach Art und Entsorgungsmethode	EN22	2
R-B10*	Gesamtgewicht des gefährlich eingestuften Abfalls und dessen Entsorgung	EN22	1
Soziales			
R-C1	Anzahl Mitarbeiter nach mindestens zwei Kategorien (Region, Geschlecht, Qualifikation)	LA1	2
R-C2	Anzahl leitende Mitarbeiter nach mindestens zwei Kategorien (Region, Geschlecht, Qualifikation)	LA13	2
R-C3	Verhältnis Grundgehalt Männer/Frauen	LA14	2
R-C4	Mitarbeiterfluktuation	LA2	2
R-C5	Gesundheit: Arbeitsunfälle, Verletzungen, Ausfalltage, Todesfälle	LA7	2
R-C6	Anteil Mitarbeiter mit Kollektivvereinbarungen/ in Gewerkschaften	LA4	1
R-C7	Weiterbildungsumfang: zeitlicher oder finanzieller Aufwand	LA10	2
R-C8	Gesamtzahl der Vorfälle von Diskriminierung	HR4	1
R-C9	Verhinderung von Kinderarbeit: Umfang und Ergebnis von Audits	HR6	1
R-C10	Verhinderung von Zwangs- und Pflichtarbeit: Umfang und Ergebnis von Audits	HR7	1
Summe			32
Branchenspezifische Leistungsindikatoren (ersetzen die mit * gekennzeichneten Leistungsindikatoren)			
Banken/Versicherungen			
R-D1	Vermögen in socially responsible investments (SRI) nach Kriterien	FS8/11	2
R-D2	Umfang Mikrofinanzierung/Mikroversicherung: Anzahl Kunden, investierte Mittel, Prämien	FS13/14	2
Handel			
R-D1	Kältemittel: Angaben zu Fluorkohlenwasserstoffen (FKW, H-FKW)	-	2
R-D2	Anteil fair gehandelter Produkte/Bioproducte am Umsatz oder Sortiment	-	2
Transport			
R-D1	Flottenzusammensetzung im Hinblick auf CO ₂ -Effizienz	LT2	2
R-D2	Nutzung erneuerbarer Energiequellen zum Transport: Anteil Bio-Treibstoffe, Anzahl Hybrid-Fahrzeuge	LT3/4	2

Abkürzung	Leistungsindikator	Bezug zu G3.1	Faktor
	Telekommunikation		
R-D1	Angaben zum Schutz von Kundendaten: Anzahl Verstösse, Anzahl und Ergebnis von Audits	PR8	2
R-D2	Anzahl der allein und gemeinsam genutzten Sendemasten	IO8	2
	Software		
R-D1	Angaben zum Schutz von Kundendaten: Anzahl Verstösse, Anzahl und Ergebnis von Audits	PR8	2
R-D2	Angebote für gemeinnützige Organisationen: Anzahl oder Wert vergünstigter Software, Umfang konkreter Initiativen	-	2
	Personaldienstleister		
-	Erhöhung der Gewichtungsfaktoren bei C4, C5, C6 und C7 um jeweils 1	-	-

Tabelle 2: Überprüfte Elemente der Nachhaltigkeitsberichterstattung mit Gewichtung

Sämtliche Informationselemente sind gemäss Vorgaben der GRI quantitativ zu erfassen. Die GRI gibt im Rahmen von Indikatorprotokollen ausserdem detaillierte Hinweise zur Berechnung der einzelnen Kennzahlen. Es wird geprüft, ob das jeweilige Informationselement berichtet wird (Umfang der Nachhaltigkeitsberichterstattung) und falls eine Berichterstattung vorliegt, ob diese quantitativ oder verbal ausfällt (Form der Nachhaltigkeitsberichterstattung). Als *verbale* Berichterstattung gelten Erläuterungen und Erörterungen sowie quantitative Informationen, aus denen die zu berichtende Kennzahl nicht ohne weitere Annahmen berechnet werden kann.⁴ Eine *quantitative* Berichterstattung impliziert die numerische Angabe der jeweiligen Kennzahl für das Berichterstattungsjahr für das Gesamtunternehmen. Auch wenn im Rahmen der Indikatorprotokolle zu den einzelnen Leistungsindikatoren mitunter mehrere zu berichtende Kennzahlen aufgelistet sind, wird im Rahmen dieses Messkonzepts bereits die Berichterstattung einer dieser Kennzahlen für das jeweilige Informationselement als „quantitative“ Berichterstattung bewertet (vgl. exemplarisch R-C5 mit Bezug zu LA7). Die gewichteten Bewertungen der Informationselemente werden zu einer Gesamtbewertung aufsummiert, wobei sowohl der Umfang als auch das quantitative/verbale Ausmass der Nachhaltigkeitsberichterstattung jeweils auf einer Skala von 0 (niedrigste Ausprägung) bis 32 (höchste Ausprägung, entspricht Summe aller Gewichtungsfaktoren) gemessen wird.

Das dargestellte Messkonzept ist geeignet, Umfang (1. Analysestufe) und quantitatives/verbales Ausmass (2. Analysestufe) der Nachhaltigkeitsberichterstattung zu bestimmen und damit die zentralen Forschungsfragen dieser Untersuchung zu beantworten. Da die Unterscheidung der Informationsart in quantitativ/verbal an der Bewertung und nicht am Informationselement anknüpft, können die Probleme der Messkonzepte bisheriger Studien weitestgehend überwunden werden. Allerdings bedingt die Fokussierung auf quantitative Leistungsindikatoren eine überschaubare Anzahl an Indikatoren, die jedoch vergleichbar zu bisheriger Studien ausfällt (mit Ausnahme der Untersuchung von Clarkson et al. 2008).

⁴ Beispiele hierfür sind die numerische Berichterstattung für lediglich ein ausgewähltes Produkt oder einen ausgewählten Teilbereich des Unternehmens oder die ausschliessliche Angabe der prozentualen Veränderung des Leistungsindikators gegenüber dem Vorjahr.

Durch eine detaillierte Beschreibung der Leistungsindikatoren und Bewertungsweise können Objektivität und Reliabilität der Messung sichergestellt werden. Auch die Validität des Messansatzes erscheint im Hinblick auf die Zielsetzung erfüllt. Dem möglichen Einwand, wonach das vorgestellte Messkonzept keine vollständige Bewertung der Nachhaltigkeitsberichterstattung wie bei anderen Studien erlaubt (vgl. exemplarisch *Quick/Knocinski* 2006; *Morhardt et al.* 2002), muss entgegengestellt werden, dass nicht alle denkbaren Berichtselemente sinnvoll in quantitativ und verbal getrennt werden können. Dementsprechend erscheint die Konzentration auf ausgewählte, relevante Leistungsindikatoren der Nachhaltigkeitsberichterstattung gerechtfertigt. Durch die Berücksichtigung der zwei Dimensionen Ökologie und Soziales wird weiterhin dem Kriterium der Mehrdimensionalität Rechnung getragen. Der folgende Abschnitt zeigt, dass auch die Anforderung der Kongruenz der beiden Messkonzepte erfüllt ist.

3.3 Messung der Nachhaltigkeitsperformance

In Anlehnung an eine Definition von Performance im Management-Kontext (vgl. *Breuer et al.* 2010, 2328) wird die *Nachhaltigkeitsperformance* eines Unternehmens definiert als seine tatsächliche Leistung im Hinblick auf ökonomische, ökologische und soziale Aspekte. Für die vorliegende Untersuchung erfolgt aus den in Abschnitt 3.1 genannten Gründen eine Konzentration auf die ökologischen und sozialen Aspekte der Nachhaltigkeitsperformance. Die Beurteilung der Performance erfordert die Messbarkeit sowie Vergleichbarkeit mittels geeigneter Bezugspunkte. Aufgrund der aufgezeigten Schwächen bisheriger Messansätze sind diese als Grundlage für die Entwicklung eines neuen Messkonzepts nicht geeignet. Vielmehr wird ein iteratives Vorgehen gewählt, dessen Ausgangspunkt entsprechend des in Abschnitt 3.1 angeführten Kriteriums der Kongruenz die Informationsitems des Messkonzepts der Nachhaltigkeitsberichterstattung darstellen.

In einem ersten Schritt wurde die Berichterstattungspraxis sämtlicher Unternehmen des Untersuchungssamples zu diesen Informationsitems analysiert. Dabei zeigte sich, dass zahlreiche der in Abschnitt 3.2 berücksichtigten Leistungsindikatoren überhaupt nicht oder nicht in quantitativer Form berichtet werden. Des Weiteren berichten die Unternehmen bezüglich eines Leistungsindikators mitunter *verschiedene* quantitativ messbare Sachverhalte (vgl. Leistungsindikator R-C5). Der vorliegende Ansatz zur Messung der Nachhaltigkeitsperformance berücksichtigt diese Problematik, indem eine Konzentration auf jene Leistungsindikatoren des Messkonzepts zur Bewertung der Nachhaltigkeitsberichterstattung erfolgt, zu denen die meisten Unternehmen des Untersuchungssamples quantitative Angaben vornehmen. Die nachfolgende Tabelle gibt einen Überblick über die Leistungsindikatoren des Messkonzepts der Nachhaltigkeitsperformance. Für jeden Leistungsindikator werden Berechnung und Masseinheit sowie der Bezug zum Messkonzept aus Abschnitt 3.2 aufgezeigt. Insgesamt werden acht Leistungsindikatoren erhoben, davon jeweils vier für die Dimensionen Ökologie und Soziales.

Abkürzung	Leistungsindikator	Berechnung	Masseneinheit	Bezug Bericht- erstattung
<i>Ökologie</i>				
P-B1	Energieverbrauch	(direkter + indirekter Energieverbrauch) / Anzahl MA	MWh / MA	R-B3
P-B2	Wasserentnahme	(gesamte Wasserentnahme – Kühlwasser) / Anzahl MA	m ³ / MA	R-B4
P-B3	Treibhausgasemissionen	(direkte + indirekte Emissionen in CO ₂ -Äquivalenten) / Anzahl MA	t / MA	R-B5
P-B4	Abfallmenge	Gesamtgewicht des Abfalls / Anzahl MA	t / MA	R-B9
<i>Soziales</i>				
P-C1	Mitarbeiterfluktuation	Anzahl aller Austritte / Anzahl MA	Prozent	R-C4
P-C2	Weiterbildung	Gesamte Weiterbildungszeit / Anzahl MA	h / MA	R-C7
P-C3	Frauenanteil in den obersten Exekutiv- und Kontrollorganen	(Anzahl Frauen in VS + AR bzw. GL + VR) / Anzahl Mitglieder bei- der Gremien	Prozent	R-C2
P-C4	Verantwortungsge- rechte Vergütung des obersten Exekutivvor- gans	(Gesamtvergütung VS bzw. GL / Anzahl der Mitglieder) / Anzahl MA	€ / MA	-

AR Aufsichtsrat GL Geschäftsleitung MA Mitarbeiter in Vollzeitäquivalenten
VR Verwaltungsrat VS Vorstand

Tabelle 3: Indikatoren zur Messung der Nachhaltigkeitsperformance

Es wird deutlich, dass gemäss dem Kriterium der Mehrdimensionalität die wesentlichen Aspekte der beiden analysierten Dimensionen der Nachhaltigkeit jeweils Berücksichtigung finden. Ausserdem zeigt ein Vergleich der Leistungsindikatoren der beiden Messkonzepte, dass auch das Kriterium der Kongruenz weitestgehend erfüllt ist. Lediglich für Indikator P-C4 gibt es keine direkte Entsprechung zur Messung der Nachhaltigkeitsberichterstattung, da aufgrund gesetzlicher Offenlegungspflichten eine Erfassung dieses Leistungsindikators im Rahmen der Nachhaltigkeitsberichterstattung nicht zweckmässig ist (vgl. Art. 663b und Art. 663c Abs. 3 OR sowie § 314 Abs. 1 Nr. 6a HGB).

Nachfolgend werden Besonderheiten der verwendeten Leistungsindikatoren kurz erläutert. Die Leistungsindikatoren P-B1 bis P-B4 stellen klassische umweltbezogene Leistungsindikatoren zu Energieverbrauch, Wasserentnahme, Treibhausgasemissionen sowie zur Abfallmenge dar. Bei sämtlichen Leistungsindikatoren erfolgt eine Berücksichtigung der Unternehmensgrösse durch Relativierung der Kennzahl über die Gesamtzahl an beschäftigten Mitarbeitern (in Vollzeitäquivalenten). Bei den Treibhausgasemissionen wird sofern vorhanden ein Abgleich mit den Angaben des Carbon Disclosure Projects vorgenommen. Es werden die tatsächlich emittierten Treibhausgase ohne Einbezug einer Kompensation durch Emissionszertifikate berücksichtigt. Die Leistungsindikatoren P-C1 bis P-C3 spie-

geln die unternehmensinterne, d.h. mitarbeiterbezogene Perspektive der sozialen Dimension wider, Leistungsindikator P-C4 die unternehmensexterne, d.h. gesellschaftliche Perspektive. Leistungsindikator P-C3, welcher einen Aspekt der Chancengleichheit von Frauen und Männern misst, berücksichtigt lediglich den Anteil von Frauen in Geschäftsleitung und Verwaltungsrat bzw. in Vorstand und Aufsichtsrat. Weitere Indikatoren, wie das Gehaltsverhältnis von Frauen und Männern auf gleicher Qualifikationsstufe bzw. der Anteil von Frauen in Leitungspositionen im gesamten Unternehmen, bleiben aufgrund fehlender oder nicht vergleichbarer Angaben der einzelnen Unternehmen unberücksichtigt. Leistungsindikator P-C4 misst die durchschnittliche Vergütung eines Geschäftsleitungs- bzw. Vorstandsmitglieds je Mitarbeiter und dient als Indikator für die gesellschaftliche Akzeptanz der Geschäftsleitungs- bzw. Vorstandsvergütung. Analog zu den Leistungsindikatoren P-B1 bis P-B4 erfolgt auch hier eine Berücksichtigung der Unternehmensgrösse durch den Einbezug der Mitarbeiterzahl im Nenner. Diese Relativierung soll dem Verantwortungsumfang der Geschäftsleitung, welcher sich auch in einer höheren Vergütung ausdrückt, Rechnung tragen. Auch wenn die absolute Höhe der durchschnittlichen Vergütung von Geschäftsleitung bzw. Vorstand je Mitarbeiter keine direkten Rückschlüsse auf ihre gesellschaftliche Akzeptanz zulässt, so kann dennoch vermutet werden, dass die gesellschaftliche Akzeptanz mit zunehmender Vergütungshöhe sinkt und diese somit als Indikator für die soziale Dimension der Nachhaltigkeitsperformance eines Unternehmens herangezogen werden kann.

Analog zu Abschnitt 3.2 werden die acht Leistungsindikatoren zu einem Gesamtwert verdichtet. In einem ersten Schritt erfolgt eine möglichst präzise und einheitliche quantitative Erfassung der einzelnen Leistungsindikatoren gemäss den Angaben zur Berechnung in *Tabelle 3*. Hierfür werden die von den Unternehmen im Rahmen der Nachhaltigkeitsberichterstattung sowie der Geschäftsberichterstattung publizierten Informationen manuell ausgewertet. Die möglichst präzise Datenerfassung ist mit zahlreichen Schwierigkeiten aufgrund von ungenauen, unvollständigen oder uneinheitlichen Angaben bei der Berichterstattung der Unternehmen verbunden. Soweit möglich werden direkt die jeweiligen Angaben des Unternehmens übernommen. Falls die Informationen lediglich für ausgewählte Bereiche, Regionen oder Mitarbeitergruppen ausgewiesen sind, werden die vorhandenen Daten auf das Gesamtunternehmen hochgerechnet. Dabei lassen sich Verzerrungen aufgrund der teilweise unvollständigen Berichterstattung nicht ausschliessen. Sofern überhaupt keine Angaben vorliegen, enthält der Datensatz an dieser Stelle fehlende Werte. Ist dies für mehr als zwei Indikatoren je Dimension der Fall, wird der Datensatz komplett ausgeschlossen. In einem zweiten Schritt werden die erhobenen Daten auf eine Skala [0, 100] transformiert. Hierfür werden die analysierten Unternehmen gemäss ihrer Branche in die Klassen Dienstleistung, Industrie und Energieversorgung eingeteilt. Insbesondere im Hinblick auf die ökologische Dimension soll hierdurch sichergestellt werden, dass die Bewertung den branchenüblichen Gegebenheiten Rechnung trägt. In jeder Klasse werden der „schlechtesten“ Ausprägung eines Leistungsindikators (z.B. beim Energieverbrauch dem höchsten Wert, bei der Weiterbildung dem niedrigsten Wert) null und der „besten“ Ausprägung 100 Bewertungspunkte zugewiesen. Die dazwischen liegenden Werte werden proportional reskaliert. Durch Bildung des arithmetischen Mittelwerts der Bewertungen der einzelnen Leistungsindikatoren erhält man die durchschnittliche Bewertung der zwei Dimensionen. Fehlende Werte werden so durch den Mittelwert der übrigen Indikatoren der jeweiligen Dimension approximiert. Das arithmetische Mittel der Bewertung der beiden Dimensionen

ergibt wiederum den Gesamtwert der Nachhaltigkeitsperformance, welcher auf einer theoretischen Skala [0, 100] liegt. Die getrennte Transformation der Originalwerte für die einzelnen Branchenklassen soll den stark unterschiedlichen Ausprägungen in den jeweiligen Klassen Rechnung tragen. Dies gilt insbesondere für die Dimension Ökologie, jedoch muss auch für die Dimension Soziales die tendenziell höhere Bedeutung des Humankapitals für Dienstleistungsunternehmen berücksichtigt werden. Die Berücksichtigung von Branchenklassen erfordert eine hinreichend grosse Anzahl an Unternehmen in den jeweiligen Klassen. Grundsätzlich ist auch eine branchenspezifischere Unterscheidung denkbar, um eine bessere Vergleichbarkeit der Daten innerhalb dieser Klassen zu gewährleisten. Allerdings muss berücksichtigt werden, dass auch bei einer engen branchenspezifischen Klasseneinteilung andersartige Geschäftsmodelle der jeweiligen Unternehmen unterschiedliche Performancedaten implizieren. Eine direkte Vergleichbarkeit der Daten wäre auch dann nicht ohne Weiteres gegeben.

Trotz dieser Einschränkungen ist das vorgestellte Messkonzept geeignet, die aufgezeigten Schwächen bisheriger Ansätze im Hinblick auf die in Abschnitt 3.1 angeführten Gütekriterien weitestgehend zu überwinden. Durch die Anknüpfung des Messkonzepts an zentrale ökologische und soziale Leistungsindikatoren kann die Objektivität der Messung sichergestellt werden. Dies wirkt sich auch positiv auf die Reliabilität der Messung aus, welche lediglich durch die Reskalierung der Daten auf Basis der minimalen und maximalen Ausprägung beeinträchtigt sein könnte. Auch die Validität ist im Vergleich zu bisherigen Messkonzepten als weitestgehend erfüllt zu beurteilen. Zwar ist eine vollständige Erfassung sämtlicher ökologischer und sozialer Auswirkungen wünschenswert, dies ist jedoch aufgrund der Berichterstattungspraxis nicht möglich. Diese Problematik wird durch eine hohe Relevanz der berücksichtigten Leistungsindikatoren gemildert. Im Hinblick auf die Reskalierung der Originaldaten muss einschränkend auf die relativ grobe Unterscheidung von Branchenklassen hingewiesen werden. Weitestgehend erfüllt sind die Kriterien Kongruenz und Mehrdimensionalität, da wesentliche Elemente der Nachhaltigkeitsberichterstattung auch bei der Performancemessung in den Bereichen Ökologie und Soziales berücksichtigt werden.

Unabhängig von der Erfüllung der Gütekriterien ist das vorgestellte Messkonzept lediglich im Hinblick auf die Anlehnung an unternehmensseitig zur Verfügung gestellte Daten und eine damit einhergehende mangelnde Überprüfbarkeit kritisch zu beurteilen. Eine Verzerrung der Ergebnisse zur Nachhaltigkeitsperformance durch fehlende oder falsche Angaben der berichtenden Unternehmen ist nicht auszuschliessen. Dieser Effekt wird durch den Ausschluss von Unternehmen, welche keine Informationen zur Nachhaltigkeitsperformance ihres Unternehmens veröffentlichen, verstärkt. Im Hinblick auf Unternehmen, die einen oder zwei fehlende Werte je Dimension aufweisen und daher noch berücksichtigt werden, ist die Verzerrung für die Hypothesenprüfung jedoch unkritisch. Sofern Unternehmen mit schlechter Performance tatsächlich mehr fehlende Werte aufweisen, wird die Performance durch die Mittelwertbildung über vorhandene Daten tendenziell überschätzt. Dies schwächt die im vierten Abschnitt ermittelten Zusammenhänge ab, es bewirkt jedoch gerade keine Verzerrung in Richtung der Bestätigung unserer Hypothesen. Zukünftig könnte die Problematik infolge einer zunehmenden Verbreitung der externen Überprüfung (Assurance) von im Rahmen der Nachhaltigkeitsberichterstattung publizierten Informationen (vgl. *Kolk* 2010) an Bedeutung verlieren. Insgesamt stellt das Messkonzept daher im Vergleich zu den Ansätzen bisheriger Studien eine deutliche Verbesserung dar.

4. Empirische Ergebnisse

4.1 Bestimmung des Untersuchungssamples

Die vorliegende Studie basiert auf einer Analyse der Nachhaltigkeitsberichterstattung von 50 Deutschen und Schweizer Grossunternehmen. Neben reinen Nachhaltigkeitsberichten werden auch Geschäfts-, Umwelt- und Personalberichte sowie andere themenbezogene Online-Publikationen einbezogen. Diese Vorgehensweise berücksichtigt, dass noch nicht alle Unternehmen ihre Nachhaltigkeitsberichterstattung in einem Dokument bündeln. Die Grundgesamtheit umfasst sämtliche Unternehmen, welche zum 31.12.2011 im Deutschen Aktienindex (DAX) oder im Swiss Leader Index (SLI) kotiert sind. Von der Analyse ausgeschlossen werden Unternehmen der Energieversorgungsbranche, da die Stichprobe lediglich zwei Unternehmen dieser Branche (E.ON und RWE) beinhaltet. Diese Klassengrösse ist für eine aussagekräftige Reskalierung der erhobenen Daten zur Nachhaltigkeitsperformance nicht ausreichend. Ergänzend wird die Nachhaltigkeitsberichterstattung und -performance der Detailhändler Migros und Coop analysiert, da diese im Hinblick auf die öffentliche Wahrnehmung eine besondere Bedeutung in der Schweiz innehaben. Der Analysezeitraum umfasst das Berichterstattungsjahr 2010. Für zehn Unternehmen (Actelion, Nobel Biocare, Fresenius Medical Care, Fresenius SE, Infineon, Logitech, Swatch, Synthes, Transocean sowie Zurich) liegen keine Informationen zur Nachhaltigkeit in auswertbarem Umfang vor, weshalb sich die Grösse der Stichprobe auf 50 Unternehmen reduziert. Von diesen können 19 Unternehmen der Dienstleistungsbranche (inklusive Handel) und 31 Unternehmen der Industriebranche (inklusive Transport) zugeordnet werden.

Die Konzentration auf Deutsche und Schweizer Grossunternehmen trägt mehreren Überlegungen Rechnung. So zeigt eine Reihe von empirischen Untersuchungen einen signifikant positiven Einfluss der Unternehmensgrösse auf den Umfang bzw. die Qualität der Nachhaltigkeitsberichterstattung (vgl. *Clarkson et al.* 2011; *Clarkson et al.* 2008; *Patten* 2002; *Bewley/Li* 2000). Dieser Effekt dürfte bei einem Untersuchungssample mit hinreichend grossen Unternehmen weniger stark ausfallen. Zudem ist hierdurch sichergestellt, dass unternehmensintern ausreichend Ressourcen zur Erstellung eines Nachhaltigkeitsberichts und zur aktiven Steuerung der Nachhaltigkeitsperformance zur Verfügung stehen. Des Weiteren erzielen Grossunternehmen in der Regel einen relativ hohen Umsatz und beschäftigen eine grosse Anzahl an Mitarbeitern, so dass sie daran gemessen einen hohen Anteil der gesamten Unternehmenslandschaft abbilden. Da es bislang keine empirische Untersuchung des Zusammenhangs zwischen Nachhaltigkeitsperformance und -berichterstattung für europäische Unternehmen gibt, erscheint eine Konzentration auf Deutsche und Schweizer Grossunternehmen in einer ersten Studie gerechtfertigt. Letztlich führt die aufwändige Form der Datenerhebung zu einer Begrenzung des Stichprobenumfangs. Sämtliche Daten zur Nachhaltigkeitsberichterstattung und -performance wurden von den Verfassern manuell erhoben.

4.2 Deskriptive Ergebnisse

Die deskriptiven Ergebnisse zu Umfang und Form der Nachhaltigkeitsberichterstattung sind in *Tabelle 4* dargestellt. Für jeden Leistungsindikator des Messkonzepts – mit Ausnahme der branchenspezifischen Leistungsindikatoren – ist der Anteil an Unternehmen getrennt nach Industrie- und Dienstleistungsbranche dargestellt, der diesen nicht, verbal oder quantitativ berichtet.

	Industrieunternehmen				Dienstleistungsunternehmen			
	<i>n</i>	<i>nicht berichtet</i>	<i>verbal berichtet</i>	<i>quantitativ berichtet</i>	<i>n</i>	<i>nicht berichtet</i>	<i>verbal berichtet</i>	<i>quantitativ berichtet</i>
R-B1	31	32 %	32 %	35 %	19	26 %	32 %	42 %
R-B2	31	42 %	48 %	10 %	19	42 %	16 %	42 %
R-B3	31	0 %	26 %	74 %	19	0 %	26 %	74 %
R-B4	31	10 %	35 %	55 %	19	0 %	21 %	79 %
R-B5	31	0 %	26 %	74 %	19	0 %	11 %	89 %
R-B6*	28	25 %	14 %	61 %	-	-	-	-
R-B7	31	19 %	10 %	71 %	19	84 %	5 %	11 %
R-B8*	28	21 %	21 %	57 %	-	-	-	-
R-B9	31	6 %	29 %	65 %	19	11 %	47 %	42 %
R-B10*	28	29 %	7 %	64 %	-	-	-	-
Mittelwert		18 %	25 %	57 %		23 %	23 %	54 %
R-C1	31	0 %	19 %	81 %	19	0 %	11 %	89 %
R-C2	31	16 %	52 %	32 %	19	21 %	79 %	0 %
R-C3	31	55 %	45 %	0 %	19	47 %	47 %	5 %
R-C4	31	23 %	19 %	58 %	19	16 %	16 %	68 %
R-C5	31	10 %	16 %	74 %	19	11 %	63 %	26 %
R-C6	31	13 %	45 %	42 %	19	21 %	32 %	47 %
R-C7	31	10 %	32 %	58 %	19	16 %	37 %	47 %
R-C8	31	23 %	52 %	26 %	19	26 %	47 %	26 %
R-C9	31	6 %	84 %	10 %	19	5 %	58 %	37 %
R-C10	31	10 %	81 %	10 %	19	11 %	63 %	26 %
Mittelwert		17 %	45 %	39 %		17 %	45 %	37 %

Tabelle 4: Deskriptive Ergebnisse zu Umfang und Form der Nachhaltigkeitsberichterstattung

Die Ergebnisse zeigen, dass insbesondere die ökologischen Leistungsindikatoren von einem hohen Anteil an Unternehmen quantitativ berichtet werden. Im Durchschnitt etwas geringer fällt der Anteil an Unternehmen mit quantitativer Berichterstattung bei den sozialen Leistungsindikatoren aus. Für einige Indikatoren (R-B2, R-C3, R-C9 und R-C10) wird deutlich, dass diese von Industrieunternehmen überhaupt nicht oder überwiegend verbal berichtet werden. So wird beispielsweise von keinem Unternehmen dieser Gruppe eine konkrete quantitative Angabe zum Verhältnis des Grundgehalts zwischen Männern und Frauen (R-C3) gemacht. Obwohl diese Thematik sowohl in Deutschland als auch in der Schweiz von den Medien intensiv diskutiert wird, unterlassen die Unternehmen an dieser Stelle eine quantitative Berichterstattung. Von den Dienstleistungsunternehmen werden hingegen zu den Indikatoren R-B7, R-C2 und R-C3 nur vereinzelt quantitative Informationen berichtet. Während dies bei R-B7 (andere Luftemissionen) vermutlich der geringen Relevanz des Indikators zugeschrieben werden kann, deutet der Verzicht auf quantitative Angaben bei R-C2 (Anzahl leitende Mitarbeiter nach Kategorien) eher auf eine gezielt zurückhaltende Informationspolitik hin. Ähnlich zu den Industrieunternehmen berichten auch die beobachteten Dienstleistungsunternehmen kaum über das Verhältnis des Grundgehalts zwischen Männern und Frauen (R-C3).

In *Tabelle 5* sind die wesentlichen deskriptiven Merkmale der Leistungsindikatoren sowie der Gesamtbewertung der Nachhaltigkeitsperformance jeweils getrennt für Industrie- und Dienstleistungsunternehmen zusammengefasst.

	P-B1	P-B2	P-B3	P-B4	P-C1	P-C2	P-C3	P-C4	Score
Industrieunternehmen									
n	31	26	31	25	19	18	31	31	31
MW	211,59	940,61	119,02	5,19	0,08	17,59	0,09	61,53	71,16
MD	45,70	300,50	14,32	3,07	0,09	18,14	0,09	42,42	72,39
SD	363,07	2.078,77	275,01	5,82	0,05	6,57	0,06	55,65	13,40
MIN	3,60	8,54	1,75	0,15	0,01	2,02	0	4,08	41,74
MAX	1.583,33	10.498,00	1.263,85	22,18	0,22	30,3	0,21	197,59	90,73
Dienstleistungsunternehmen									
n	19	19	19	18	17	14	19	19	19
MW	16,23	43,82	4,39	0,72	0,12	23,64	0,11	137,37	64,41
MD	14,78	18,90	3,98	0,30	0,11	18,00	0,09	57,87	63,84
SD	9,30	69,15	2,85	0,99	0,05	18,23	0,06	206,41	11,99
MIN	3,15	10,06	1,58	0,06	0,05	6,40	0,04	7,29	39,24
MAX	32,81	309,09	12,80	3,38	0,25	80,00	0,25	724,32	83,84

MW Mittelwert MD Median SD Standardabweichung MIN Minimum MAX Maximum

Tabelle 5: Deskriptive Ergebnisse zur Nachhaltigkeitsperformance

Die Ergebnisse zeigen teilweise deutliche Unterschiede bei den Leistungsindikatoren für Industrie- und Dienstleistungsunternehmen. Während die ökologischen Leistungsindikatoren wie erwartet bei Industrieunternehmen im Durchschnitt schlechter ausfallen als bei Dienstleistungsunternehmen, erhält man für die soziale Dimension ein weniger eindeutiges Ergebnis. Hier schneiden Industrieunternehmen bei der Mitarbeiterfluktuation (P-C1) sowie bei der verantwortungsgerechten Vergütung des obersten Exekutivorgans (P-C4) besser und bei den übrigen beiden Leistungsindikatoren tendenziell schlechter ab als Dienstleistungsunternehmen. Insgesamt erscheint die getrennte Reskalierung der Leistungsindikatoren daher insbesondere im Hinblick auf die ökologische Dimension zweckmässig, um die unterschiedlichen Streuungen der Performancedaten bei Industrie- und Dienstleistungsunternehmen zu berücksichtigen. Allerdings weisen das deutliche Auseinanderfallen von arithmetischem Mittelwert und Median sowie die grosse Streuung der Ergebnisse bei den Industrieunternehmen darauf hin, dass für diese Gruppe eine weitere Unterteilung angebracht sein könnte.

4.3 Hypothesenprüfung

Für die Hypothesenprüfung wird auf Pearson-Korrelationsanalysen zurückgegriffen.⁵ Eine Überprüfung von *Hypothese H₁* ergibt einen positiven und signifikanten Zusammenhang ($r=0,33$; $p=0,02$; $n=50$) zwischen der Nachhaltigkeitsperformance und dem Umfang der

⁵ Die Ergebnisse des Shapiro-Wilk-Tests auf Normalverteilung indizieren mit einer Ausnahme für sämtliche zu untersuchenden Variablen keine Ablehnung der Normalverteilungshypothese bei einem Signifikanzniveau von 10%.

Nachhaltigkeitsberichterstattung. In *Abbildung 1* ist dieser Zusammenhang – unterschieden nach Industrie- und Dienstleistungsunternehmen – grafisch dargestellt.

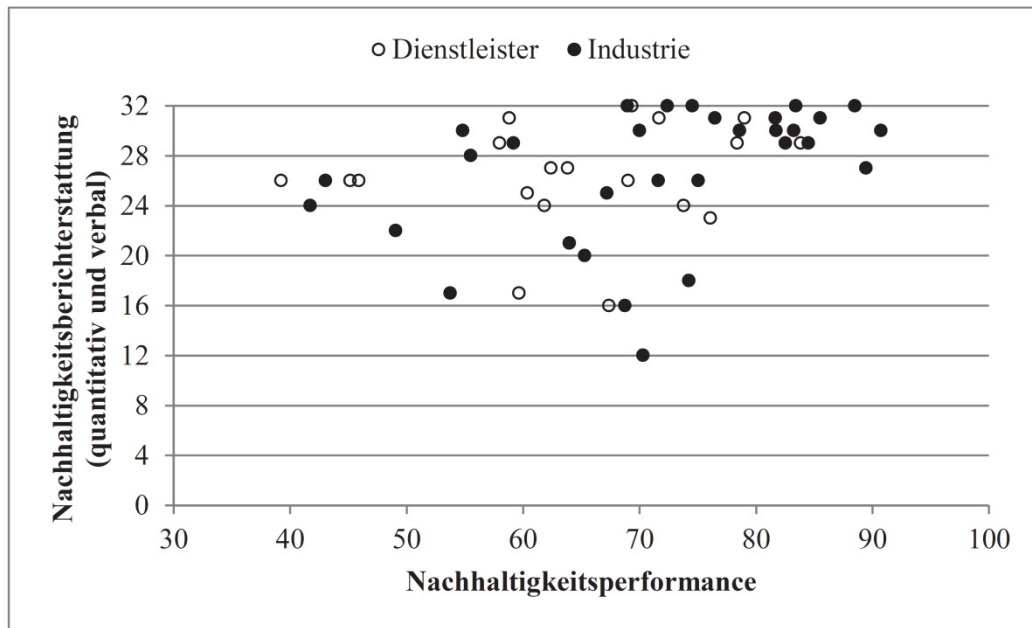


Abbildung 1: Grafische Zusammenhangsanalyse zu Hypothese H₁

Auch die Ergebnisse zur quantitativen Form der Nachhaltigkeitsberichterstattung stehen im Einklang mit unserer Argumentation. Der Zusammenhang zwischen Nachhaltigkeitsperformance und dem quantitativen Ausmass der Nachhaltigkeitsberichterstattung ist positiv und signifikant ($r=0,46$; $p=0,00$; $n=50$). Eine grafische Veranschaulichung mittels Streupunktdiagramm findet sich in *Abbildung 2*.

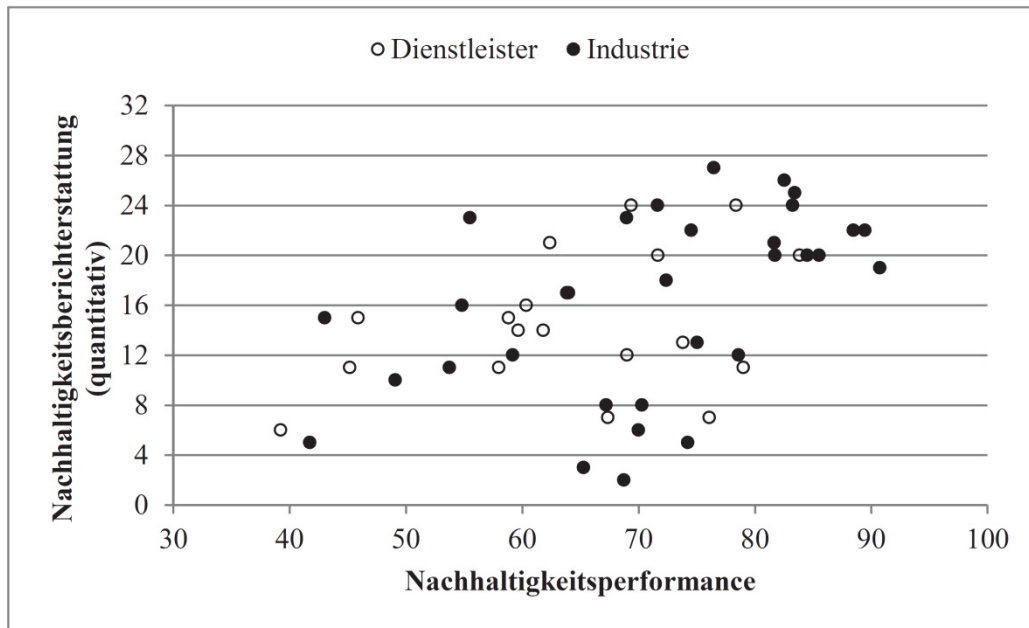


Abbildung 2: Grafische Zusammenhangsanalyse zu Hypothese H_{2a}

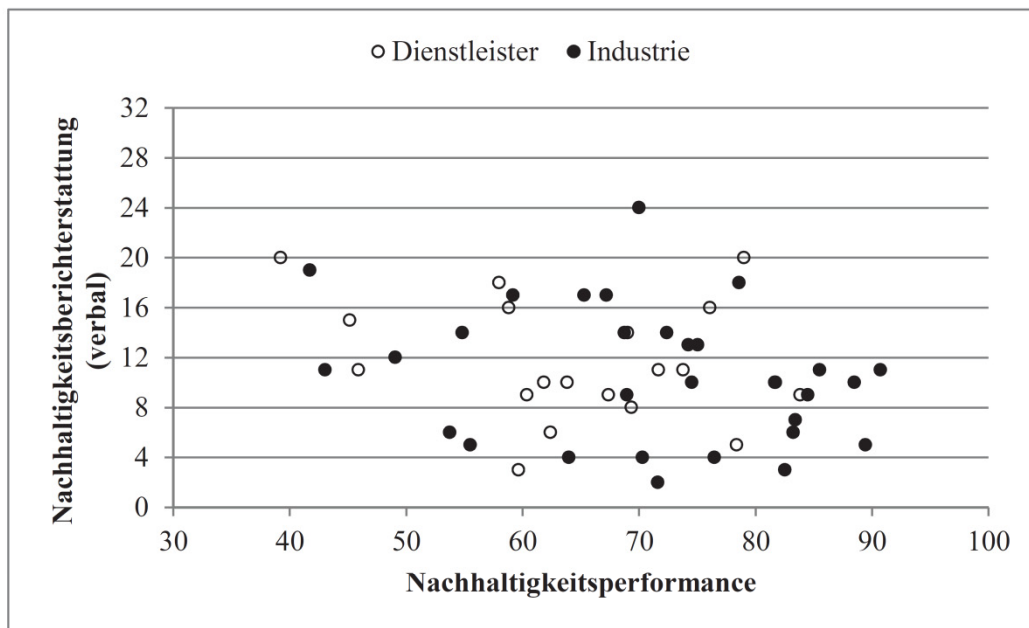


Abbildung 3: Grafische Zusammenhangsanalyse zu Hypothese H_{2b}

Je besser die Nachhaltigkeitsperformance, desto eher veröffentlichen Unternehmen quantitative Informationen im Rahmen der Nachhaltigkeitsberichterstattung. Beide Hypothesen, welche einen Teil der Nachhaltigkeitsberichterstattung auf Basis der VDT erklären, werden somit gestützt.

Im Rahmen von *Hypothese H_{2b}* wird der Argumentation der LT folgend ein negativer Zusammenhang zwischen Nachhaltigkeitsperformance und verbaler Form der Nachhaltigkeitsberichterstattung postuliert. Auch diese Hypothese wird durch die empirischen Ergebnisse gestützt. Es besteht ein negativer und auf dem 10%-Niveau signifikanter Zusammenhang ($r=-0,27$; $p=0,06$; $n=50$) zwischen der Nachhaltigkeitsperformance und dem Ausmass der verbalen Nachhaltigkeitsberichterstattung (vgl. *Abbildung 3*).

Insgesamt bestätigen somit die empirischen Ergebnisse sämtliche Hypothesen der Untersuchung. Nachfolgend soll nun mittels Sensitivitätsanalysen die Robustheit der Ergebnisse untersucht werden.

Hierfür wird in einem ersten Schritt der Ansatz zur Messung der Nachhaltigkeitsberichterstattung im Hinblick auf die Wahl der Gewichtungsfaktoren analysiert. Es wird untersucht, inwiefern die Gewichtung der einzelnen Berichtselemente die Untersuchungsergebnisse beeinflusst. Neu fliessen sämtliche Berichtselemente mit einem Gewichtungsfaktor von 1 in das alternative Bewertungskonzept ein.⁶ In diesem alternativen Modell fallen die Ergebnisse hinsichtlich Richtung und Signifikanz nahezu identisch zu unserem Ausgangsmodell aus (H_1 : $r=0,33$; $p=0,02$; $n=50$. H_{2a} : $r=0,47$; $p=0,00$; $n=50$. H_{2b} : $r=-0,26$; $p=0,07$; $n=50$). Somit wird deutlich, dass unsere Untersuchungsergebnisse nicht durch die Wahl der Gewichtung getrieben werden.

In einem zweiten Schritt soll der Ansatz zur Messung der Nachhaltigkeitsperformance untersucht werden. In Anlehnung an die Vorgehensweise bisheriger Studien (vgl. *Choi/Patten 2007*) wird für die statistischen Auswertungen auf eine binäre Skalierung der Nachhaltigkeitsperformance zurückgegriffen. Hierfür wird das intervallskalierte Mass der Nachhaltigkeitsperformance durch eine Unterscheidung in high/low performer (über/unter Median) ersetzt. Die Ergebnisse eines Mittelwertvergleichs stützen erneut die *Hypothesen H₁* und *H_{2a}*, jedoch fällt das Ergebnis für *Hypothese H_{2b}* nicht signifikant aus.⁷ Die Analyse zeigt somit, dass eine intervallskalierte Messung der Nachhaltigkeitsperformance einer binären Unterscheidung überlegen ist. Des Weiteren wird die Problematik fehlender Werte bei einzelnen Leistungsindikatoren untersucht. In unserem Grundmodell werden diese fehlenden Werte durch den Mittelwert der übrigen Indikatoren in der jeweiligen Dimension ersetzt. Diese Vorgehensweise könnte eine Abschwächung der untersuchten Effekte bewirken, da gemäss VDT anzunehmen ist, dass vor allem schlechte Performancewerte nicht berichtet werden. Andererseits kann insbesondere für die ökologische Dimension von einem ähnlichen Niveau der Leistungsindikatoren innerhalb eines Unternehmens ausgegangen werden, so dass der Mittelwert der übrigen Performancewerte eine zumindest akzeptable Approximation des tatsächlichen Wertes darstellen könnte. Sämtliche Analysen werden erneut ausschliesslich für vollständige Datensätze durchgeführt. Die Ergebnisse fallen bezüglich Richtung und Signifikanz vergleichbar zu unserem Ausgangsmodell aus (H_1 : $r=0,38$; $p=0,07$; $n=24$. H_{2a} : $r=0,56$; $p=0,00$; $n=24$. H_{2b} : $r=-0,34$; $p=0,11$; $n=24$). Lediglich die Ergebnisse für *Hypothese H_{2b}* sind in diesem Modell knapp nicht signifikant zum 10%-Niveau. Hierbei ist jedoch der mit $n=24$ geringere Stichprobenumfang zu beachten. Inge-

6 Davon ausgenommen sind die mit * gekennzeichneten Leistungsindikatoren, bei welchen die Gewichtung unverändert erhalten bleibt. Die Substitution mit branchenspezifischen Leistungsindikatoren würde andernfalls zu unterschiedlichen maximal möglichen Gesamtbewertungen für die einzelnen Unternehmen führen.

7 Die Ergebnisse lauten: H_1 : MW(low)=24,88; MW(high)=28,20; $p(T<t)=0,01$. H_{2a} : MW(low)=13,12; MW(high)=17,96; $p(T<t)=0,00$. H_{2b} : MW(low)=11,76; MW(high)=10,24; $p(T>t)=0,15$.

samt scheint daher der Umgang mit fehlenden Werten in unserem Messkonzept der Nachhaltigkeitsperformance zu keiner systematischen Verzerrung der Ergebnisse zu führen.

In einem letzten Schritt wird die Robustheit der Ergebnisse bei Berücksichtigung von Kontrollvariablen untersucht. Den Ergebnissen bisheriger Untersuchungen folgend wird für die Grösse der untersuchten Unternehmen kontrolliert (vgl. *Clarkson et al.* 2011; *Clarkson et al.* 2008; *Patten* 2002; *Bewley/Li* 2000). Die Unternehmensgrösse wird über die Mitarbeiterzahl operationalisiert. Für die Analyse erfolgt basierend auf dem Median eine Unterscheidung in kleine und grosse Unternehmen. Die Hypothesen werden erneut getrennt für diese beiden Samples überprüft. Wiederum wird die Richtung der Zusammenhänge für sämtliche Hypothesen bestätigt. Die Zusammenhänge sind für grosse Unternehmen durchgängig signifikant, für kleine Unternehmen sind lediglich die Ergebnisse zu Hypothese H_{2a} signifikant.⁸ Somit scheinen insbesondere grosse Unternehmen eine systematische Nachhaltigkeitsberichterstattung gemäss der theoretischen Erklärungsansätze vorzunehmen. Dies könnte beispielsweise auf die stärkere Wahrnehmung grosser Unternehmen durch das gesellschaftliche Umfeld und Investoren zurückzuführen sein. Darüber hinaus wird die Berichterstattung für Deutsche und Schweizer Unternehmen getrennt analysiert, da bisherige empirische Studien auf Unterschiede in der Nachhaltigkeitsberichterstattung zwischen verschiedenen Ländern hinweisen (vgl. exemplarisch *Hummel* 2012, 178). Auch für diese Variante wird die Richtung des Zusammenhangs für sämtliche Hypothesen bestätigt. Die Zusammenhänge sind jedoch durchgängig lediglich für Deutsche und nicht für Schweizer Unternehmen statistisch signifikant.⁹ Eine mögliche Erklärung könnte die LT liefern, wonach das gesellschaftliche Umfeld in der Schweiz gegenüber den Unternehmen weniger kritisch als in Deutschland eingestellt sein könnte. Dies würde mit einem geringeren Legitimitätsdruck auf die Unternehmen in der Schweiz einhergehen.¹⁰

Insgesamt erweist sich der methodische Ansatz zur Messung von Nachhaltigkeitsberichterstattung und Nachhaltigkeitsperformance somit für die Untersuchung der Fragestellungen als geeignet. Auch unsere Argumentation, wonach sowohl VDT als auch LT jeweils einen Teil der Nachhaltigkeitsberichterstattung von Unternehmen erklären können, wird durch die Ergebnisse der empirischen Überprüfung sowie die Sensitivitätsanalysen gestützt.

5. Fazit und kritische Diskussion

Ausgangspunkt des vorliegenden Beitrags ist der Zusammenhang zwischen der Nachhaltigkeitsperformance eines Unternehmens und Umfang und Form seiner Nachhaltigkeitsberichterstattung. Mögliche Erklärungsansätze werden im Rahmen der Hypothesenformulie-

8 Die Ergebnisse lauten: H_1 : $r=0,23$; $p=0,28$; $n=25$ für kleine Unternehmen bzw. $r=0,45$; $p=0,02$; $n=25$ für grosse Unternehmen. H_{2a} : $r=0,34$; $p=0,09$; $n=25$ für kleine Unternehmen bzw. $r=0,55$; $p=0,00$; $n=25$ für grosse Unternehmen. H_{2b} : $r=-0,18$; $p=0,40$; $n=25$ für kleine Unternehmen bzw. $r=-0,41$; $p=0,04$; $n=25$ für grosse Unternehmen. Alternativ haben wir die Unternehmensgrösse über Umsatz und Marktkapitalisierung (jeweils zum Ende 2010) operationalisiert und erhalten bezüglich Richtung und Signifikanz nahezu identische Ergebnisse.

9 Die Ergebnisse lauten: H_1 : $r=0,11$; $p=0,59$; $n=25$ für Schweizer Unternehmen bzw. $r=0,53$; $p=0,01$; $n=25$ für Deutsche Unternehmen. H_{2a} : $r=0,31$; $p=0,14$; $n=25$ für Schweizer Unternehmen bzw. $r=0,55$; $p=0,00$; $n=25$ für Deutsche Unternehmen. H_{2b} : $r=-0,30$; $p=0,15$; $n=25$ für Schweizer Unternehmen bzw. $r=-0,36$; $p=0,08$; $n=25$ für Deutsche Unternehmen.

10 Die jeweiligen Subsamples auf Basis von Unternehmensgrösse und Landeszugehörigkeit sind heterogen, so dass eine Zuordnung der Effekte zu den beiden Kontrollvariablen plausibel ist. Innerhalb der Gruppe der grossen Unternehmen beträgt der Anteil Schweizer Unternehmen 36% gegenüber einem Anteil von 50% in der Gesamtstichprobe.

rung auf Basis der LT und der VDT abgeleitet. Angesichts methodischer Schwächen bisheriger Studien bei der Operationalisierung der zentralen Untersuchungsvariablen steht im Mittelpunkt des Beitrags die Entwicklung neuer Konzepte zur Messung von Umfang und Form der Nachhaltigkeitsberichterstattung sowie der Nachhaltigkeitsperformance. Erstmals wird bei der Bewertung der berichteten Informationen durchgängig zwischen quantitativer und verbaler Nachhaltigkeitsberichterstattung unterschieden. Das vorgestellte Messkonzept zur Bestimmung der Nachhaltigkeitsperformance ermöglicht eine objektive und mehrdimensionale Erfassung wie sie nach Kenntnis der Verfasser zum aktuellen Zeitpunkt in dieser Form und in diesem Zusammenhang bislang nicht vorgenommen wurde. Beide Messkonzepte werden hinsichtlich der Erfüllung zentraler Gütekriterien kritisch untersucht und auf eine Stichprobe von 50 Deutschen und Schweizer Grossunternehmen angewandt. Die Ergebnisse zeigen, dass die Messkonzepte für eine Überprüfung der theoretischen Erklärungsansätze geeignet sind. Zudem wird im Unterschied zu bisherigen Untersuchungen deutlich, dass beide Theorien jeweils einen *Teil* der Nachhaltigkeitsberichterstattung erklären können.

Einschränkend muss darauf hingewiesen werden, dass für die statistischen Auswertungen aufgrund der relativ geringen Stichprobengrösse der Einsatz multivariater Analysemethoden nicht zweckmässig ist. Hier bieten sich Anhaltspunkte für weitere Forschungsarbeiten. Neben einer Vergrösserung der Datenbasis ist insbesondere die Berücksichtigung einer Reihe von Kontrollvariablen für eine abschliessende Analyse der Zusammenhänge erforderlich. So weisen die Ergebnisse der vorliegenden Untersuchung auf grundlegende Unterschiede zwischen Deutschen und Schweizer Unternehmen hin. Durch eine Ausweitung der Datenbasis auf weitere Länder könnte der Einfluss länderspezifischer Besonderheiten im Rahmen zukünftiger Untersuchungen näher betrachtet werden. Zudem könnte die Erfassung der Nachhaltigkeitsperformance weiter ausdifferenziert werden. Denkbar ist beispielsweise eine branchenspezifischere Beurteilung anstelle der einfachen Unterscheidung in Industrie- und Dienstleistungsunternehmen.

Letztlich beruhen die Daten zur Messung der Nachhaltigkeitsperformance auf Angaben der berichtenden Unternehmen. Die Ausführungen in diesem Beitrag zeigen, dass auch bei Anwendung der GRI-Richtlinien zur Nachhaltigkeitsberichterstattung eine Vergleichbarkeit der berichteten Daten nicht ohne Weiteres gegeben ist. Dadurch wird die Auswahl der Indikatoren für das Messkonzept stark limitiert. Hier könnten einheitliche und bindende Vorschriften zum Inhalt und zur Prüfung der Nachhaltigkeitsberichterstattung deren Transparenz und Qualität deutlich erhöhen. So würde die Erfassung der Nachhaltigkeitsperformance und ein darauf basierender Vergleich verschiedener Unternehmen einfacher.

Literaturhinweise

- Al-Tuwaijri, S.A./Christensen, T.E./Hughes, K.E.* (2004): The relations among environmental disclosure, environmental performance, and economic performance: a simultaneous equations approach, in: *Accounting, Organizations and Society*, Jg. 29, Nr. 5, S. 447–471.
- Ashforth, B.E./Gibbs, B.W.* (1990): The double edge of legitimization, in: *Organization Science*, Jg. 1, Nr. 2, S. 177–194.
- Bewley, K./Li, Y.* (2000): Disclosure of environmental information by Canadian manufacturing companies: A voluntary disclosure perspective, in: *Advances in Environmental Accounting & Management*, Jg. 1, S. 201–226.

- Bortz, J./Döring, N. (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, 4. Auflage, Heidelberg.
- Breuer, W./Breuer, C./Heldt, C./Lackes, R./Siepermann, M. (2010): Performance, in: Gabler Wirtschaftslexikon, 17. Auflage, Wiesbaden, S. 2328.
- Chatterji, A.K./Levine, D.I./Toffel, M.W. (2009): How Well Do Social Ratings Actually Measure Corporate Social Responsibility?, in: Journal of Economics & Management Strategy, Jg. 18, Nr. 1, S. 125–169.
- Cho, C.H./Guidry, R.P./Hageman, A.M./Patten, D.M. (2012): Do actions speak louder than words? An empirical investigation of corporate environmental reputation, in: Accounting, Organizations and Society, Jg. 37, Nr. 1, S. 14–25.
- Cho, C.H./Patten, D.M. (2007): The role of environmental disclosures as tools of legitimacy: A research note, in: Accounting, Organizations and Society, Jg. 32, Nr. 7, S. 639–647.
- Cho, C.H./Patten, D.M./Roberts, R.W. (2006): Corporate Political Strategy: An Examination of the Relation between Political Expenditures, Environmental Performance, and Environmental Disclosure, in: Journal of Business Ethics, Jg. 67, Nr. 2, S. 139–154.
- Clarkson, P.M./Li, Y./Richardson, G.D./Vasvari, F.P. (2008): Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis, in: Accounting, Organizations and Society, Jg. 33, Nr. 4, S. 303–327.
- Clarkson, P.M./Overell, M.B./Chapple, L. (2011): Environmental Reporting and its Relation to Corporate Environmental Performance, in: Abacus, Jg. 47, Nr. 1, S. 27–60.
- Davis, K. (1973): The case for and against business assumption of social responsibilities, in: Academy of Management Journal, Jg. 16, Nr. 2, S. 312–322.
- Dawkins, C./Fraas, J.W. (2011): Coming Clean: The Impact of Environmental Performance and Visibility on Corporate Climate Change Disclosure, in: Journal of Business Ethics, Jg. 100, Nr. 2, S. 303–322.
- de Villiers, C./van Staden, C.J. (2006): Can less environmental disclosure have a legitimizing effect? Evidence from Africa, in: Accounting, Organizations and Society, Jg. 31, Nr. 8, S. 763–781.
- Deegan, C. (2002): Introduction: The legitimising effect of social and environmental disclosures – a theoretical foundation, in: Accounting, Auditing & Accountability Journal, Jg. 15, Nr. 3, S. 282–311.
- Dowling, J./Pfeffer, J. (1975): Organisational legitimacy: social values and organisational behavior, in: Pacific Sociological Review, Jg. 18, Nr. 1, S. 122–136.
- Elkington, J. (1999): Cannibals with Forks: The Triple Bottom Line of 21st Century Business, Oxford.
- Fekrat, M.A./Inclan, C./Petroni, D. (1996): Corporate Environmental Disclosures: Competitive Disclosure Hypothesis Using 1991 Annual Report Data, in: The International Journal of Accounting, Jg. 31, Nr. 2, S. 175–195.
- Global Reporting Initiative (2012): What is GRI?, unter <https://www.globalreporting.org/information/about-gri/what-is-GRI/Pages/default.aspx> (Abfragedatum 1.11.2012).
- Gray, R./Kouhy, R./Lavers, S. (1995): Corporate social and environmental reporting: A review of the literature and a longitudinal study of UK disclosure, in: Accounting, Auditing & Accountability Journal, Jg. 8, Nr. 2, S. 47–77.
- Healy, P.M./Palepu, K.G. (2001): Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature, in: Journal of Accounting and Economics, Jg. 31, Nr. 1–3, S. 405–440.

- Hoffmann, T. (2011): Unternehmerische Nachhaltigkeitsberichterstattung. Eine Analyse des GRI G3.1-Berichtsrahmens, Lohmar.
- Hughes, S.B./Anderson, A./Golden, S. (2001): Corporate environmental disclosures: are they useful in determining environmental performance?, in: Journal of Accounting and Public Policy, Jg. 20, Nr. 3, S. 217–240.
- Hummel, K. (2012): Internationaler Vergleich der Nachhaltigkeitsberichterstattung, in: Controlling, Jg. 24, Nr. 3, S. 175–180.
- Ingram, R.W./Frazier, K.B. (1980): Environmental Performance and Corporate Disclosure, in: Journal of Accounting Research, Jg. 18, Nr. 2, S. 614–622.
- Kolk, A. (2010): Determinants of the Adoption of Sustainability Assurance Statements: An International Investigation, in: Business Strategy and the Environment, Jg. 19, Nr. 3, S. 182–198.
- Li, Y./Richardson, G.D./Thornton, D. (1997): Corporate Disclosure of Environmental Liability Information: Theory and Evidence, in: Contemporary Accounting Research, Jg. 14, Nr. 3, S. 435–474.
- Milne, M.J./Patten, D.M. (2002): Securing organizational legitimacy: An experimental decision case examining the impact of environmental disclosures, in: Accounting, Auditing & Accountability Journal, Jg. 15, Nr. 3, S. 372–405.
- Morhardt, J.E./Baird, S./Freeman, K. (2002): Scoring Corporate Environmental and Sustainability Reports Using GRI 2000, ISO 14031 and other Criteria, in: Corporate Social Responsibility and Environmental Management, Jg. 9, Nr. 4, S. 215–233.
- O'Donovan, G. (2002): Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory, in: Accounting, Auditing & Accountability Journal, Jg. 15, Nr. 3, S. 344–371.
- Patten, D.M. (2002): The relation between environmental performance and environmental disclosure: a research note, in: Accounting, Organizations and Society, Jg. 27, Nr. 8, S. 763–773.
- Quick, R./Knocinski, M. (2006): Nachhaltigkeitsberichterstattung, in: Zeitschrift für Betriebswirtschaft, Jg. 76, Nr. 6, S. 615–650.
- Rockness, J.W. (1985): An Assessment of the Relationship between US Corporate Environmental Performance and Disclosure, in: Journal of Business Finance & Accounting, Jg. 12, Nr. 3, S. 339–354.
- Sethi, S.P. (1978): Advocacy Advertising – The American Experience, in: California Management Review, Jg. 21, Nr. 1, S. 55–68.
- Suchman, M.C. (1995): Managing Legitimacy: Strategic and Institutional Approaches, in: The Academy of Management Review, Jg. 20, Nr. 3, S. 571–610.
- van der Laan Smith, J./Adhikari, A./Tondkar, R.H. (2005): Exploring differences in social disclosures internationally: A stakeholder perspective, in: Journal of Accounting and Public Policy, Jg. 24, Nr. 2, S. 123–151.
- Verrecchia, R.E. (1983): Discretionary Disclosure, in: Journal of Accounting and Economics, Jg. 5, Nr. 1, S. 179–194.
- Wiseman, J. (1982): An Evaluation of Environmental Disclosures Made in Corporate Annual Reports, in: Accounting, Organizations and Society, Jg. 7, Nr. 1, S. 53–63.

Katrin Hummel, Dr. rer. pol., ist Oberassistentin am Lehrstuhl für Accounting, insbesondere Unternehmensrechnung und Controlling am Institut für Betriebswirtschaftslehre der Universität Zürich.

Christian Schlick, Dipl.-Kfm., ist Assistent am Lehrstuhl für Accounting, insbesondere Unternehmensrechnung und Controlling am Institut für Betriebswirtschaftslehre der Universität Zürich.

Anschrift: Universität Zürich, Institut für Betriebswirtschaftslehre, Lehrstuhl für Accounting, insbes. Unternehmensrechnung und Controlling, Plattenstr. 14, CH-8032 Zürich, Tel.: +41 (0) 44/634-29-83, Fax: +41 (0) 44/634-49-12, E-Mails: katrin.hummel@business.uzh.ch; christian.schlick@business.uzh.ch

The Relationship between Sustainability Performance and Sustainability Disclosure Quality – Hard Numbers Beat Smooth Talk¹

ABSTRACT

This study investigates the relationship between corporate sustainability performance (CSP) and the quality of corporate sustainability disclosure (CSD). We respond to the call in recent research to change the research setting to investigate the overlap between voluntary disclosure theory (VDT) and legitimacy theory (LT) more rigorously. By concentrating exclusively on the quality – and not the quantity – of CSD, we are able to show the simultaneous applicability of predictions derived from both VDT and LT. With respect to CSD, we derive and assess information items by distinguishing between higher and lower quality disclosures. Our approach to CSP measurement is based on multi-dimensional proxies for a firm's environmental and social performance. The results from a sample of 195 European companies illuminate how VDT and LT overlap. We find evidence that poor sustainability performers use lower quality CSD to protect their sustainability image and that higher quality CSD is associated with superior CSP.

INTRODUCTION

In recent decades, the concept of corporate sustainability has received growing attention from many segments of society. The sizable numbers of ethically managed mutual funds and ethically screened portfolios illustrate that it may not suffice for companies to be merely financially successful; instead, firms may also be expected to act responsibly with respect to their environment. This development has been accompanied by various sustainability initiatives, such as the Carbon Disclosure Project, the Principles for Responsible Investment and the Equator Principles. Investors increasingly demand information about companies' sustainability performances. The recent action of 24 institutional investors that complained about the "lack of information as a result of limited or non-disclosure of ESG [i.e., environmental, social and corporate governance] data" (Abberley, quoted in Reuters, 2011) emphasizes the importance of high quality sustainability disclosures in the capital markets. In an open letter, these same institutional investors asked the world's 30 largest stock exchanges to overhaul current listing rules to "force companies to improve their sustainability reporting" (Reuters, 2011). The information provided by a company must enable investors and other stakeholders to assess the company's true sustainability performance.

However, the relationship between corporate sustainability performance (CSP) and corporate sustainability disclosure (CSD) remains ambiguous both theoretically and empirically. In essence, there are two theoretical concepts involved. According to legitimacy theory (LT), on the one hand, companies use CSD to improve the public perception of their CSP (Deegan, 2002). Researchers therefore interpret a negative relationship between CSP and the level of

¹ This paper is co-authored by Dr. Katrin Hummel, University of Zurich.

CSD as an indication of the applicability of LT (Dawkins & Fraas, 2011; Patten, 2002). On the other hand, based on voluntary disclosure theory (VDT) it is argued that a company with good CSP is incentivized to disclose information regarding its performance to increase its market value. Following this stream of research, a positive relationship between CSP and the level of CSD is posited. Thus, opposing predictions regarding the relationship between CSP and CSD are typically derived from these two theories and the mixed empirical results from prior studies cannot clarify this relationship. Recent research has therefore asked whether these two theories are partially overlapping rather than mutually exclusive and has found some preliminary evidence to justify this reasoning (Clarkson, Li, Richardson, & Vasvari, 2008). Building on this research, our study provides a more rigorous examination of this relationship by concentrating on the quality dimension of CSD.

With respect to the measurement of both CSD and CSP, prior research has not yet established a consistent approach. For instance, some studies concentrate on the reporting of a few disclosure items (Patten, 2002), whereas others attempt to comprehensively capture CSD (Clarkson et al., 2008). With respect to scoring systems, most studies employ a binary scoring (disclosure vs. non-disclosure), whereas some utilize ordinal ratings (Al-Tuwaijri, Christensen, & Hughes, 2004). This heterogeneity is even more remarkable with respect to the measurement and operationalization of CSP. Some studies rely exclusively on a single indicator of environmental performance, such as emissions or waste (Clarkson, Overell, & Chapple, 2011), whereas other studies' approaches are based on rating metrics provided by specialized rating agencies, such as Kinder, Lydenberg, Domini (KLD, today MSCI) (Cho & Patten, 2007; Cho, Patten, & Roberts, 2006; Dawkins & Fraas, 2011). Because the validity and/or reliability of these measurement approaches are questionable, several researchers have called for the development of more refined measures of CSP (Hong & Andersen, 2011).

Against this background, the aim of this study is to investigate the relationship between CSP and the quality of CSD in a refined setting and thereby to further illuminate how VDT and LT overlap. In so doing, our study makes several contributions to the literature. First, we respond to the call by Clarkson et al. (2008) for future research and switch the focus of enquiry from CSD quantity to CSD quality. Second, we present improved and transparent measurement approaches for CSP and CSD quality that include the social dimension of corporate sustainability. Third, we apply these measurement schemes to a sample of 195 publicly traded European companies. Empirical findings from our study support the preliminary evidence of previous research on the overlap between VDT and LT. In addition, the results from our study enhance the knowledge about other determinants of CSD quality, which is not sufficiently understood in Western Europe (Fifka, 2013).

The remainder of this paper is structured as follows. The next section reviews the related literature and develops the hypotheses. The section following that explains our research design and focuses, in particular, on the measurement schemes for CSD quality and CSP that are applied in this study. In addition, this section describes the data sample and the empirical model. Next, we provide descriptive results and our findings from regression analyses and robustness checks. The final section concludes the paper.

LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

From a theoretical perspective, nearly all previous empirical studies in this field are based on VDT or LT (Al-Tuwaijri et al., 2004; Cho, Guidry, Hageman, & Patten, 2012; Cho & Patten, 2007; Clarkson et al., 2008; Clarkson et al., 2011; de Villiers & van Staden, 2006; Patten, 2002). These studies focus primarily on the *level* of CSD (i.e., CSD quantity) by classifying each disclosure item as either disclosed (equal to 1) or non-disclosed (equal to 0). A positive relationship between CSP and the level of CSD is then interpreted as support for the applicability of VDT, and a negative relationship is attributed to LT. Such a research design neglects the potential overlap between these two theories because the theories are treated as incompatible with one another. The mixed results of previous empirical studies (for a positive relationship between CSP and CSD see Al-Tuwaijri et al., 2004; Clarkson et al., 2008; for a negative relationship see Cho and Patten, 2007; de Villiers and van Staden, 2006) are characteristic of this research setting. Moreover, these empirical findings are seldom unambiguous. For instance, Clarkson et al. (2008) interpret a positive relationship between environmental performance and the level of environmental disclosure as evidence for the application of VDT but refer to LT to explain “interesting patterns in the data” (Clarkson et al., 2008, p. 323). Attempting to investigate the overlap between the two theories (see also Deegan, 2002; Clarkson et al., 2011) more rigorously, researchers have developed measurement schemes for CSD that distinguish between “monetary” and “non-monetary” (Patten, 2002) or “hard” and “soft” disclosure items (Clarkson et al., 2008). These researchers argue that VDT predicts a positive relationship between CSP and the use of “hard” disclosures, whereas LT predicts a negative relationship between CSP and the use of “soft” disclosures. However, because firms can score high on the disclosure of both “hard” and “soft” items, the results do not permit a clear distinction between the applicability of the two theories. For instance, a positive relationship between environmental performance and the level of hard disclosure can be attributed to either VDT (i.e., better environmental performance is related to more hard disclosure) or LT (i.e., worse environmental performance is related to less hard disclosure). Similarly, a negative relationship between environmental performance and soft disclosure can support both VDT (better environmental performance is related to less soft disclosure) and LT (worse environmental performance is related to more soft disclosure). Accordingly, Clarkson et al. (2008, p. 325) call for a switch in the “focus of enquiry”. To be able to clearly distinguish between the applicability of VDT and the applicability of LT, we must examine the *quality* of CSD in lieu of the *quantity* of CSD. Thus, we develop a research design in which both theories can be tested simultaneously and not exclusive of one another, which may thereby provide theoretical and empirical insights into how these theories may overlap.

VDT explains why managers of a risky asset voluntarily disclose information about this asset by formulating an equilibrium threshold level of disclosure in which the cost associated with the disclosure is precisely equal to the information’s effect on the price of the risky asset (Verrecchia, 1983). The model offers an explanation for why rational managers withhold unfavorable information below a critical threshold level of disclosure. Although this theory originally referred to the voluntary disclosure of *financial* information, researchers have also applied it to explain the voluntary disclosure of *non-financial* information (Bewley & Li, 2000;

Clarkson et al., 2008; Li, Richardson, & Thornton, 1997). They argue that a company with superior CSP voluntarily discloses non-financial information to reveal its true performance type and potentially increase its market value (Clarkson et al., 2008). VDT's reasoning builds on decision-useful information, i.e., information that is objective, reliable, comparable and relevant to investors. Following previous research (Al-Tuwaijri et al., 2004; Bewley & Li, 2000; Cho & Patten, 2007; Clarkson et al., 2008), we distinguish between quantitative (*higher* quality) and qualitative (*lower* quality) information. We define quantitative disclosure as the disclosure of relevant and comparable numerical data that fulfills or exceeds precisely defined quality requirements. Qualitative disclosure refers to any lower quality disclosures that can be either numeric or non-numeric but that do not fulfill these criteria. We argue that the reasoning of VDT applies primarily to quantitative disclosure because quantitative disclosure is more reliable and comparable than qualitative disclosure (Al-Tuwaijri et al., 2004; Clarkson et al., 2008; Clarkson et al., 2011). Companies with superior CSP disclose higher quality quantitative information to signal their true (unobservable) performance type. Moreover, the disclosure of quantitative information cannot easily be mimicked by companies with poor CSP (Clarkson et al., 2008), which enables superior performers to distinguish themselves from poorly performing companies. Following this reasoning, we expect to find that companies with superior CSP primarily disclose quantitative (higher quality) information. We formally state our hypothesis as follows:

H1: There is a positive relationship between CSP and the level of quantitative (higher quality) CSD.

Legitimacy theory (LT) offers another theoretical explanation for the voluntary disclosure of non-financial information. Suchman (1995, p. 574) defines legitimacy as “a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions”. This abstract concept of society is more precisely delineated by Freeman's (1984, p. 25) definition of stakeholders as “groups [...] that can affect, or are affected by, the accomplishment of organizational purpose” (Wood, 1991). According to Dowling and Pfeffer (1975, p. 122), any disparity between the actions of an organization and the socially constructed system of norms is a “threat to organizational legitimacy” and therefore to the organization's long-term survival (Davis, 1973). Sethi (1978) offers three legitimation tactics to close such legitimacy gaps between business performance and societal expectations. He recommends changing public perceptions first through information and second through symbolic actions; if these actions prove ineffective, he recommends altering the business performance itself. Similarly, O'Donovan (2002) describes four possible response tactics to legitimacy threats: avoidance, altering social values, shaping societal perceptions of the organization and compliance with societal expectations.

A wide-ranging lack of sustainability can threaten the legitimacy of a company. LT suggests that particularly poor performing companies use CSD as a legitimation tactic to influence public perceptions of their CSP (Deegan, 2002). In this case, CSD serves as a tool to build or maintain a company's sustainability image. Thus, poor performing companies prefer *qualitative*

(lower quality) information (e.g., information that is superficial, unnecessary, incomplete, incomparable, ambiguous or non-numeric) because stakeholders can easily identify *quantitative* (higher quality) information as poor performance values that could be interpreted as a lack of sustainability. Although there are professional stakeholders that cannot be easily misled by such lower quality information, the general public's perception of a firm's sustainability is also influenced by non-professional stakeholders that may not notice such disclosure tactics. Consequently, we expect that companies with poor CSP prefer qualitative (lower quality) to quantitative (higher quality) reporting on critical issues to close an existing or impending legitimacy gap. We posit the following relationship:

H2: There is a negative relationship between CSP and the level of qualitative (lower quality) CSD.

The two hypotheses delineated above are intended to clearly distinguish between quantitative (higher quality) and qualitative (lower quality) CSD. Consistent with both VDT and LT, companies may also opt for non-disclosure on a specific issue. Thus, evidence falsifying one of the hypotheses does not necessarily allow inferences for the other hypothesis. Additionally, our arguments are based on the assumption that firms' reporting behavior is driven by incentives to increase market value (VDT) *and at the same time* by attempts to avoid the negative consequences of threatened legitimacy (LT). These combined forces result in varying CSD quality that depends on actual CSP. Thus, the two stated hypotheses are tested simultaneously and not against one another.

RESEARCH DESIGN

Methodological requirements

To ensure the rigor of our research design, CSD and CSP must be reliably and validly measured (Krippendorff, 2013). In addition to these traditional measurement criteria, further requirements must be fulfilled to account for our specific research design. These criteria consist of congruence in the content of the CSD and CSP measurement schemes, the multi-dimensionality of the measurement schemes and informational relevance. Content-based congruence refers to the direct linkage between CSP and CSD quality measurements. To precisely assess the theoretical implications derived from VDT and LT, proxies for both CSD and CSP must capture similar content, for instance, reporting elements on emissions and corresponding emissions performance data. Otherwise changes in CSP are not reflected in the observed CSD quality. The multi-dimensionality criterion stems from the understanding of sustainability developed by Elkington (1997), as described by the triple-bottom-line approach. Both measurement schemes should cover the environmental and social dimensions of sustainability and account for different aspects within each dimension. The criterion of relevance derives from VDT and refers to the decision usefulness of information. It is also required by the

Global Reporting Initiative (GRI) G3.1 reporting guidelines (“materiality”, GRI, 2011a) and financial reporting standards (FASB, 2010).²

CSD quality measurement scheme

The GRI guidelines define sustainability reporting as “the practice of measuring, disclosing, and being accountable to internal and external stakeholders for organizational performance toward the goal of sustainable development” (GRI, 2011a, p. 3). In our study, we concentrate on a company’s voluntary disclosure of environmental and social, in particular employee-related information.³ In the field of voluntary, non-financial disclosure, content analysis has become the primary research methodology. Researchers first identify relevant information items and then assess the respective disclosures for each item.

An overview of the measurement approaches applied in previous empirical research into the relationship between CSD and CSP is provided in the appendix. Many of these studies distinguish between “monetary” or “hard” and “non-monetary” or “soft” disclosure items (Clarkson et al., 2008; Patten, 2002). Hard disclosure items focus on a firm’s disclosures related to environmental performance indicators such as data on emissions, water use and recycling, whereas soft disclosure items concentrate on the disclosure of a firm’s vision, environmental strategy and commitment to environmentally responsible management. Because the disclosure items are observed on a binary basis (disclosure vs. non-disclosure), the overall disclosure score refers to the *level* (i.e., quantity) but not to the *quality* dimension (i.e., the shades of reporting between disclosure and non-disclosure) of CSD. In this case, an empirically testable distinction between VDT and LT is difficult.⁴ For instance, Clarkson et al. (2008) report a positive and significant relationship between environmental performance and environmental disclosure for all three disclosure categories (total, hard and soft disclosures). They interpret this finding as evidence for the application of VDT but refer to LT to explain additional “interesting patterns in the data” (Clarkson et al., 2008, p. 323). Other measurement schemes are based on ordinal ratings and thus attempt to capture disclosure quality directly. In this case, a higher rating is typically assigned to quantitative disclosure, and a lower rating is assigned to non-quantitative disclosure (Aerts & Cormier, 2009; Aerts, Cormier, & Magnan, 2008; Al-Tuwaijri et al., 2004; Wiseman, 1982). Because the overall disclosure score consists of a mixture of quantitative and qualitative ratings for each item, both VDT and LT predict a positive relationship between CSP and CSD (Clarkson et al., 2011). Again, it is difficult to clearly distinguish between the applicability of each theory.

² In May 2013, the G4 reporting standard was introduced. Because companies may begin applying it in reporting year 2013 at the earliest, the paper refers to reporting standard G3.1. In addition, the G4 standard is even more focused on materiality than the G3.1 standard (GRI, 2013).

³ We do not include the economic dimension of sustainability because the reporting of key economic information is mandatory and subject to national and international accounting standards. Moreover, the investigation of various aspects of mandatory and voluntary financial disclosure represents its own line of research (for an overview, see Healy and Palepu, 2001).

⁴ A positive relationship between environmental performance and hard disclosure can support both VDT (better environmental performance is related to more hard disclosure) and LT (worse environmental performance is related to less hard disclosure). Similarly, a negative relationship between environmental performance and soft disclosure can support both VDT (better environmental performance is related to less soft disclosure) and LT (worse environmental performance is related to more soft disclosure).

Against this background, we provide a measurement scheme for CSD that concentrates on the *quality* instead of the quantity of CSD. Our index of relevant disclosure items is closely linked to the performance indicators required by GRI guidelines version 3.1. We focus on disclosure items that are required to be disclosed numerically. For each disclosure item, the GRI guidelines provide a precise description of all information that must be disclosed by a company. Each sustainability category – environmental and social – consists of seven performance indicators that must be disclosed according to GRI G3.1. These are classified by the GRI as core indicators that are generally applicable and relevant for most companies (GRI, 2011a). For each of the respective disclosure items, exactly one point is awarded for quantitative (higher quality) disclosure, qualitative (lower quality) disclosure or non-disclosure. We define quantitative (higher quality) disclosure as the disclosure of numerical data on a company-wide level that fulfill or exceed the requirements derived from GRI G3.1 and described in Table 1. If these requirements are not fulfilled, but any type of quantitative, qualitative or mixed information is presented, one point is awarded for qualitative (lower quality) disclosure. When there is no information at all the item is classified as not reported. We include sector-specific adjustments for the quantitative (higher quality) requirements in the environmental category to account for industry-specific variations in the relevance of specific disclosure items. An example is presented in the footnote.⁵ Quantitative (qualitative) disclosure scores are calculated as the sum of all quantitative (qualitative) disclosure item scores and range between zero (minimum) and 14 (maximum).

Because the differentiation between quantitative, qualitative and non-disclosure refers to the same disclosure items, our measurement scheme for CSD quality enables an empirically testable distinction between the applicability of VDT and LT. The reliability of our measure is supported by detailed descriptions of the disclosure items and measurement procedures. Moreover, the definition of each disclosure item is also provided by GRI G3.1 and is available to the disclosing companies. Concentrating on relevant disclosure items strengthens the validity of our measurement scheme. Thus, our measurement scheme enables the measurement of a company's preference for quantitative (higher quality) versus qualitative (lower quality) disclosure, but it does not provide a full assessment of all areas of CSD (i.e., the quantity of CSD). By including an environmental and social category of CSD, we account for the criterion of multi-dimensionality. The congruence criterion is discussed in the next section along with the measurement scheme for CSP.

⁵ For example, reporting element R-E7 in Table 1 requires all companies except those belonging to the industry groups 13 to 15 to report their total waste weight by type and by disposal method. A chemical company may report the following: "Our total amount of waste is 100,000 tons that are split up into 30,000 tons of hazardous waste and 70,000 tons of non-hazardous waste. 10% of our waste is recycled and 90% is sent to landfills." This statement satisfies the minimum criteria for R-E7 and is therefore coded as quantitative (higher quality) disclosure. If only the first sentence were reported it would be coded as qualitative (lower quality) disclosure, although some quantitative information is given. For the same indicator R-E7 an insurance company (industry group 14) is only required to report the total amount of its waste without details about waste types and disposal methods to be coded as quantitative (higher quality).

Table 1: Measurement scheme for corporate sustainability disclosure (CSD) quality

Code	Disclosure item	Requirements for quantitative (higher quality) score	GRI G3.1
<i>Environmental dimension</i>			
R-E1	materials used	all substantial input materials by weight or volume ^a	EN1
R-E2	energy consumption and renewables	direct and indirect energy consumption, share of renewable energy sources ^b	EN3/4
R-E3	water withdrawal	total withdrawal by source ^c	EN8
R-E4	greenhouse gas emissions	total direct and indirect emissions (GHG protocol scopes 1, 2, and 3)	EN16/17
R-E5	ozone-depleting substances and other air emissions	total emissions of ozone-depleting substances; other significant air emissions by type and weight for at least one substance; alternatively, an explicit statement of irrelevance for both ^d	EN19/20
R-E6	water discharge	total discharge by quality (emissions to water by type and weight for at least one substance; alternatively, an explicit statement of irrelevance) and destination ^e	EN21
R-E7	waste	total weight by type and disposal method ^f <i>maximum environmental score is 7</i>	EN22
<i>Social dimension</i>			
R-S1	workforce	total workforce based on at least three criteria (division, region, employment type, employment contract, qualification, age or gender)	LA1
R-S2	employee turnover	total number of employees leaving for any reason (not for a single reason only)	LA2
R-S3	collective bargaining agreements	percentage of total workforce covered by collective bargaining agreements	LA4
R-S4	safety and health	work safety and health based on at least two criteria (rates of injury, occupational diseases, lost days, absenteeism, fatalities)	LA7
R-S5	training	training (or related) time	LA10
R-S6	discrimination	total number of incidents or explicit statement that no incidents occurred (no indirect paraphrasing or references to codes of conduct)	HR4
R-S7	child, forced, and compulsory labor	scope and numerical results of audits (within company or supply chain) regarding at least one aspect <i>maximum social score is 7</i>	HR6/7

^a Adjustments for industry groups 13 to 15 (see description in Table 3): use of paper is sufficient.

^b Adjustments for industry group 12 (see description in Table 3): share of renewable energy sources is excluded. Adjustments for industry group 18 (see description in Table 3): share of renewable energy *produced*.

^c Adjustments for industry groups 4 to 6 and 13 to 15 (see description in Table 3): by source is excluded.

^d Adjustments for industry groups 4 to 6 and 13 to 15 (see description in Table 3): ozone-depleting substances *or* other significant air emissions.

^e Adjustments for industry groups 4 to 6 and 13 to 15 (see description in Table 3): by quality and destination is excluded.

^f Adjustments for industry groups 13 to 15 (see description in Table 3): by type and disposal method is excluded.

Table 1 presents the indicators for the assessment of quantitative (higher quality) and qualitative (lower quality) CSD with respect to the environmental and social reporting dimension. The first two columns contain the respective disclosure items. Column three presents requirements that are defined for quantitative (higher quality) disclosure. The fourth column indicates the link to GRI G3.1 (GRI, 2011b, 2011c, 2011d).

CSP measurement scheme

The appendix provides measurement approaches for CSP that have been applied in previous empirical studies on the relationship between CSD and CSP. Some studies are based on one or two indicators – such as emissions or waste – as proxies for overall environmental perfor-

mance (Al-Tuwaijri et al., 2004; Clarkson et al., 2008). Such an approach is likely to miss the required level of validity. Other studies rely on a wide range of different emission indicators from publically available registers. For instance, Horváthová's (2012) data stem from the European Pollutant Release and Transfer Register (E-PRTR), which is limited to industrial facilities. In this case, emissions outside a specific region (e.g., Europe) are not covered and the validity of this measurement approach with respect to international companies is thus questionable. There are also studies that utilize rating metrics provided by external agencies (Cho & Patten, 2007; Cho et al., 2006; Dawkins & Fraas, 2011). One of the most widely used ratings is the KLD (today, the MSCI) rating which is based on binary data for 14 items, in which the respective environmental performance of a company is assessed in terms of "strengths" and "weaknesses" (Chatterji, Levine, & Toffel, 2009). Researchers typically derive a total score as a measure of environmental performance by subtracting the total weaknesses from the total strengths. Although this approach is not constrained by focusing exclusively on a single environmental performance indicator, the ratings process is not fully transparent. In particular, the criteria and threshold levels for the binary assessment are not revealed. Therefore, the reliability cannot be evaluated from an outsider's perspective. Moreover, the binary assessment insufficiently accounts for substantial variances in the underlying data. As a consequence, researchers have repeatedly questioned the appropriateness of the KLD database for purposes of academic research (Chin, Hambrick, & Treviño, 2013; Chiu & Sharfman, 2011) and have called for the development of an improved measure of CSP (Hong & Andersen, 2011).

We heed this call and develop a more refined measure of CSP. To ensure content-based congruence between the CSD quality and CSP measurement schemes, our indicators for CSP are derived from the disclosure items of CSD. If there are several performance indicators for a single disclosure item (e.g., R-S4), we focus on the most important in terms of environmental and social performance. Moreover, the selection of our indicators accounts for the comparability and availability of the respective data within and across industry groups. Table 2 provides an overview of the performance indicators, the measurements of these indicators, the units of measurement and the corresponding CSD items. The measurement scheme consists of four environmental and four social performance indicators. Each indicator is precisely measured as defined in Table 2. If available, we directly refer to the data provided by the respective company. If data are provided only for certain countries, business areas or employee groups but cover at least 80% of total sales or total employees, we extrapolate the data to the entire corporation. If data are unreported, the respective performance indicator for the company contains a missing value, and companies with more than two missing values in one dimension are excluded from the sample. Moreover, performance indicators can contain valid values even when the corresponding disclosure items are of lower quality and coded as qualitative.⁶

⁶ For example, a chemical company may exclusively report its total waste weight and be coded as qualitative (lower quality) for reporting element R-E7 because further information about the types of waste and disposal methods is missing. Nevertheless, the information requirement for the corresponding performance indicator P-E4 (total weight of waste) is met.

Table 2: Measurement scheme for corporate sustainability performance (CSP)

Code	Performance indicator	Measurement	Unit	CSD link
<i>Environmental dimension</i>				
P-E1	energy consumption	(direct + indirect energy consumption) / number of employees ^a	MWh / employee	R-E2
P-E2	water withdrawal	(total water withdrawal – cooling water) / number of employees ^a	m ³ / employee	R-E3
P-E3	greenhouse gas emissions	(GHG protocol scope 1 + scope 2 emissions) / number of employees ^a	t / employee	R-E4
P-E4	total weight of waste	total weight of waste / number of employees ^a	t / employee	R-E7
<i>Social dimension</i>				
P-S1	employee turnover	total number of employees ^a who leave / number of employees ^a * 100	percent	R-S2
P-S2	lost time incident rate	number of incidents resulting in lost time from work / (total hours worked / 200,000)	incidents / h	R-S4
P-S3	employee training	total training time / number of employees ^a	h / employee	R-S5
P-S4	share of women in the highest corporate bodies ²	total number of women in the highest corporate bodies ^b / total number of members of the highest corporate bodies ^b * 100	percent	R-S6

^a If available, full-time equivalents; headcount otherwise.

^b Management board and supervisory board for two-tier system countries; board of directors for one-tier system countries.

Table 2 presents the indicators for the measurement of CSP with respect to the environmental and social dimensions. Columns one and two contain indicator names and brief descriptions. Columns three and four present formulas for calculations and the units of measurement. The last column indicates how CSP measurement is linked to CSD quality measurement in Table 1.

The original data for each indicator are arranged by industry groups and then winsorized within each industry group at the top and bottom tails at a 10% level to limit the influence of outliers (Tukey, 1962).⁷ Next, all values are transformed into a continuous [0, 1] scale per industry group by assigning “0” to the worst and “1” to the best performance indicator values and by rescaling all other values proportionally. Thus, this step allows all performance indicators with their differing measurement units to be aggregated into a total performance score. A minimum of five companies per industry group is required to define reasonable peer groups. Rescaling on an industry group basis simplifies the comparison of performance scores within these groups and improves the distribution characteristics of the CSP variable.⁸ Our final CSP

⁷ Due to fundamental differences in the data gathering procedure of our study compared with previous studies, we expect a higher likelihood of spurious outliers for our performance indicators. Such spurious outliers may result from the conversion and extrapolation of raw data (undertaken by both the reporting firms themselves and the authors).

⁸ The natural differences in absolute environmental indicator values across industry groups may otherwise complicate the analysis of the relationship between CSP and CSD. For example, all mining companies would score very low and in a narrow bandwidth in terms of environmental performance, although some are much better than others within their class. Similarly, insurance firms would typically score very high and also in a narrow bandwidth, although some firms are clearly performing worse than others in their class. Other researchers use the rank of performance data within each industry group as a proxy for a firm’s environmental performance (Clarkson, Fang, Li, & Richardson, 2013; Clarkson, Li, & Richardson, 2004; Clarkson et al.,

score is calculated as the arithmetic mean of the means of the environmental and social performance indicators. Each performance indicator is therefore weighted equally and missing values are approximated by the average of the remaining indicators within each dimension.

Our measurement scheme contains essential and well-known performance indicators that are transparently defined and thus supports the criterion of reliability. The scheme also avoids oversimplification and largely satisfies the criterion of validity. Because the chosen performance indicators are highly relevant, we mitigate the limitation that we cannot account for all the environmental and social impacts of the sample firms. Content-based congruence between the CSD quality and CSP measurement schemes and their multi-dimensionality are ensured because all key indicators of the CSD measurement scheme are also included in the CSP measurement scheme. Moreover, the criterion of relevance is fulfilled because the performance indicators derive from the GRI guidelines, which are widely considered the most commonly used international sustainability reporting standard at present (Ballou, Heitger, & Landes, 2006; Gray, 2006; KPMG, 2011).

Sample and methodological approach

The initial sample consists of 388 companies that are included in the Bloomberg European 500 index in January 2013 and are located in France, Germany, Italy, Spain, Sweden, Switzerland or the United Kingdom. Between February and August 2013, English versions of the corporate sustainability reports, if any, the annual report and any web-based sustainability disclosures by the sample firms for reporting year 2011 were identified. The authors manually assessed the CSD of each company according to the measurement scheme and recorded the data relevant for evaluating its CSP.⁹ Of the companies initially in our sample, 151 were excluded because they had more than two missing performance values in one dimension (or both dimensions). An additional 42 companies were excluded because each industry group is required to have at least five companies to allow for meaningful comparisons with peers. Table 3 shows the final sample of 195 companies from seven countries and 19 industry groups.

2008) at the cost of inter-company variance in the respective data. Another approach is to limit the sample to certain types of industries. Several studies focus on firms in environmentally sensitive industries because environmental performance and environmental disclosure can be more important to firms operating in these industries (Cho et al., 2012; Cho & Patten, 2007; Cho et al., 2006; Clarkson et al., 2008; Clarkson et al., 2011).

⁹ The sample was divided into two groups, and each author coded one group. All uncertainties were discussed and resolved by consensus.

Table 3: Sample distribution by country and by industry group

By country	n	By industry group	n
1 France	44	1 chemicals	10
2 Germany	28	2 building materials, paper, steel	16
3 Italy	16	3 mining	11
4 Spain	27	4 advertising, entertainment, media	11
5 Sweden	13	5 telecommunication	11
6 Switzerland	14	6 computers and software	5
7 United Kingdom	53	7 car manufacturers	8
		8 retail	5
		9 food and beverages	9
		10 commercial services	7
		11 pharmaceuticals	8
		12 oil and gas	8
		13 banking and financial services	30
		14 insurance	10
		15 REITS	10
		16 aerospace and defense	8
		17 engineering and construction	11
		18 electricity	7
		19 gas, water, and electricity distribution	10
Total	195	Total	195

The relationship between CSD quality and CSP along with control variables is assessed by running the following regression models. *CSDquan* refers to the level of quantitative (high quality) CSD and *CSDqual* refers to the level of qualitative (low quality) CSD.

$$CSD_{quan} = \beta_0 + \beta_1 CSP + \beta_2 ASSURE + \beta_3 ORG + \beta_4 SIZE + \beta_5 FCF + \beta_6 LEV + \beta_7 TOBIN \\ + \sum_{i=8}^{14} \beta_i COUNTRY + \sum_{j=15}^{33} \beta_j INDUSTRY + \varepsilon$$

$$CSD_{qual} = \beta_0 + \beta_1 CSP + \beta_2 ASSURE + \beta_3 ORG + \beta_4 SIZE + \beta_5 FCF + \beta_6 LEV + \beta_7 TOBIN \\ + \sum_{i=8}^{14} \beta_i COUNTRY + \sum_{j=15}^{33} \beta_j INDUSTRY + \varepsilon$$

Control variables

We rely on previous research into the determinants of voluntary non-financial disclosure to select our control variables. For instance, Fifka (2013) provides a comprehensive overview. All the variables are summarized in Table 4 and are described in greater detail below. First, we control for a company's strategic orientation toward sustainability issues, which derives from the strategic management literature (Ullmann, 1985). We draw on external assurance (*ASSURE*) and the hierarchical level of the internal organizational units focused on corporate sustainability (*ORG*) as proxies for strategic orientation. We argue that an active strategic orientation supports the disclosure of credible sustainability-related information, and we thus expect to find a positive (negative) effect on *CSDquan* (*CSDqual*). Our variable *ASSURE* indicates whether a company's CSD is assured by an external company. Previous research has

shown that obtaining external assurance is associated with the strategic integration of sustainability initiatives (Abdel-Khalik, 1993; Knechel, Salterio, & Ballou, 2007). We assume that only companies with an active strategic orientation toward sustainability bear the extra costs of external assurance to demonstrate commitment and credibility. Our variable *ORG* is measured on a four-point rating scale as shown in Table 4. Previous researchers argue that sustainability reporting requires specific reporting processes and structures (Adams, 2002; Al-Tuwaijri et al., 2004; Ballou, Casey, Grenier, & Heitger, 2012; Ruhnke & Gabriel, 2013). The hierarchical level of sustainability-related internal organizational units signals the importance that a company attaches to sustainability (Cowen, Ferreri, & Parker, 1987) and thus indicates the intensity of a firm's strategic orientation toward sustainability.

We also control for firm size (*SIZE*) because nearly all previous empirical studies show the impact of firm size on CSD (Branco & Rodrigues, 2008; Clarkson et al., 2008; Clarkson et al., 2011; Cormier, Magnan, & van Velthoven, 2005; Dawkins & Fraas, 2011; Kolk, 2003; Patten, 2002). One explanation for this effect focuses on economies of scale with respect to information production costs (Clarkson et al., 2008). Another argument refers to firm size as a proxy for other factors, primarily public visibility (Branco & Rodrigues, 2008; Dawkins & Fraas, 2011) and the extent of monitoring by analysts (Cormier et al., 2005). In either case, a positive relationship between *SIZE* and a proxy for overall disclosure can be expected. Our measure of CSD quality is more refined. Thus, an increase in size may trigger firms to switch from non-disclosure to lower quality (*CSDqual*) or higher quality (*CSDquan*) disclosure, to switch from higher quality to lower quality disclosure (legitimacy theory and increased public visibility) or the reverse (stronger demand from analysts). Because those three effects overlap and partially cancel one another out, no sign is predicted for the relationship between *SIZE* and *CSDquan* or *CSDqual*. We measure *SIZE* as the log of the number of employees at the end of the fiscal year because market capitalization is less stable over time and the book value of total assets or sales is less comparable across industries (e.g., banking and insurance).

We use the free cash flow in million euros per employee (*FCF*) at the end of the fiscal year as a proxy for financial performance. One group of researchers (Ullmann, 1985) claims that a company's financial performance determines its financial capacity to invest in and maintain CSD. In addition, these researchers argue that only financially sound companies are able to withstand the (negative) consequences of disclosing proprietary information (Cormier & Magnan, 2003; Cormier et al., 2005). Following this reasoning, the effect of financial performance on *CSDquan* and *CSDqual* might be positive. By contrast Neu, Warsame, and Pedwell (1998) conclude that companies use environmental disclosures during unprofitable years to demonstrate long-term competitive advantages resulting from environmental investments and posit a negative relationship with financial performance. A third line of research (Patten, 1991) assumes an indifferent relationship between a company's financial performance and its social disclosures, arguing that social disclosure is primarily driven by social legitimacy rather than economic legitimacy. Against the background of these contradictory theoretical considerations and inconsistent empirical results, we do not predict a sign for the relationship between financial performance and *CSDquan* or *CSDqual*.

Table 4: Variables overview

	CSDquan predicted sign	CSDqual predicted sign	Description	Source of data
<i>CSDquan</i>			level of quantitative (higher quality) CSD for reporting year 2011, measured as described in Table 1 and in the “CSD quality measurement scheme” section	manual
<i>CSDqual</i>			level of qualitative (lower quality) CSD for reporting year 2011, measured as described in Table 1 and in the “CSD quality measurement scheme” section	manual
<i>CSP</i>	+	–	level of CSP for reporting year 2011, measured as described in Table 2 and in the “CSP measurement scheme” section	manual
<i>ASSURE</i>	+	–	external assurance, equals “1” if a company’s CSD for reporting year 2011 is (at least partly) assured by an external company and “0” otherwise	manual
<i>ORG</i>	+	–	highest hierarchical level of internal organizational units focused on corporate sustainability during the 2011 reporting year, measured on a four-point rating scale: (1) a company’s corporate sustainability activities are supervised by only a few local stand-alone units with no superordinate organizational unit on the group level, (2) there is a specialized organizational unit determining and supervising the corporate sustainability strategy for the entire group, (3) at least one member of this specialized organizational unit is a representative of the highest corporate bodies, (4) a main committee of the board of directors is dedicated to the firm’s corporate sustainability activities	manual
<i>SIZE</i>			firm size, measured as the log of total employees at the end of fiscal year 2011 (full-time equivalents if available, headcount otherwise)	manual
<i>FCF</i>			financial performance, measured as the free cash flow (cash flow from operating activities – total capital expenditures) in million euros per employee at the end of fiscal year 2011 (full-time equivalents if available, headcount otherwise)	Bloomberg
<i>LEV</i>	+	–	financial leverage (average total assets / average total common equity)	Bloomberg
<i>TOBIN</i>	+	–	Tobin’s Q, measured as (market value common equity + book value long-term debt and current liabilities) / book value total assets at the end of fiscal year 2011	Bloomberg
<i>COUNTRY</i>			country of domicile dummy variables as reported in Table 3, and the reference category is the UK	Bloomberg
<i>INDUSTRY</i>			industry group dummy variables as reported in Table 3, and the reference category is banking and financial services	Bloomberg/ manual

We also include the financial leverage of a company (*LEV*) as a proxy for the informational needs of a company’s creditors. Financial leverage is measured as average total assets divided by average total common equity. It can be assumed that the monitoring demand for information by a company’s creditors increases with its leverage (Branco & Rodrigues, 2008; Clarkson et al., 2011) and that creditors are interested in a company’s CSP because CSP indicates future potential risks related to sustainability issues. Moreover, certain researchers argue

that superior CSP is essential for a company's long-term economic success and existence (McWilliams & Siegel, 2001; Porter & Kramer, 2011; Russo & Fouts, 1997). Thus, higher-leveraged companies are more dependent on creditor demands and have greater incentive to inform creditors of their true CSP type (Roberts, 1992; Ullmann, 1985). Because quantitative (higher quality) disclosure is assumed to be more reliable and comparable, we expect to find a positive (negative) effect for financial leverage on *CSDquan* (*CSDqual*).

Next, we include Tobin's Q (*TOBIN*) in our models as another control variable. Tobin's Q is measured as a firm's market value relative to the replacement cost of its assets and is used to capture information asymmetry (Al-Tuwaijri et al., 2004; Clarkson et al., 2008; Clarkson et al., 2011; Stanny & Ely, 2008). A higher Tobin's Q reflects a greater degree of information asymmetry. Following the literature on voluntary (financial) disclosure, information asymmetry between a company's managers and outside investors is assumed to be the major source of demand for (financial) disclosure (Healy & Palepu, 2001). Consequently, managers strive to decrease information asymmetry through additional disclosure. Because investors prefer quantitative information to qualitative information, we predict a positive (negative) effect of information asymmetry on *CSDquan* (*CSDqual*).

We also include country and industry group dummies as control variables. A number of empirical studies have revealed systematic, country-specific variations in firm CSD that may arise from regulatory, cultural, and societal differences (Fekrat, Inclan, & Petroni, 1996; Kolk, 2003; Kolk, Walhain, & van de Weteringen, 2001; Orij, 2010; van der Laan Smith, Adhikari, & Tondkar, 2005). Such differences are closely linked to legitimacy theory because the definition and understanding of legitimacy vary across different societies. In this context, van der Laan Smith et al. (2005) refer to the stakeholder versus shareholder orientation of a country. However, because regulatory, cultural and societal differences interfere with one another and cannot be clearly disentangled, we include firms' countries of domicile as control variables to capture such differences.¹⁰ Additionally, industry group dummies are included to account for industry-specific effects on CSD. Such industry-specific effects have been demonstrated by a substantial number of previous empirical studies (Cho & Patten, 2007; Dawkins & Fraas, 2011; Patten, 2002; Roberts, 1992). As with firm size, some researchers argue that industry affiliation serves as a proxy for public visibility and political scrutiny (Branco & Rodrigues, 2008; Cho et al., 2012; Cho & Patten, 2007; Roberts, 1992). Other researchers refer to the general institutional context to explain industry-specific reporting patterns (Cormier et al.,

¹⁰ We examine the country-specific regulatory background for our sample firms for 2011 with respect to the voluntary character of our disclosure items and find no evidence that our CSD measurement scheme is significantly biased through national regulations. We identify three countries with detailed regulations on sustainability reporting: France, Sweden and the United Kingdom. In the United Kingdom, carbon reporting has been mandatory since 2013 (Companies Act 2006 Regulations 2013). In Sweden, sustainability reporting in accordance with GRI guidelines is only required for state-owned companies, and such companies are not included in our sample (Guidelines for external reporting by state-owned companies, 2007). The reporting on some of our disclosure items is mandatory for firms in 2011 only in France (Decree 2002-221 and the Order of April 30, 2002; further requirements by Articles 75 and 225 of Grenelle II came into force in 2012). However, these French regulations do not necessarily entail quantitative disclosure as defined by our CSD measurement scheme. In addition, some European companies are obligated to report emissions data to the central European Pollutant Release and Transfer Register (E-PRTR). Nevertheless, the required data do not cover facilities outside the European Union.

2005; Patten, 2002) and argue that firms tend to compare themselves with intra-industry peers and adjust their reporting behavior accordingly. In addition, both the level of discretion and the level of data sensitivity can vary significantly across industry groups. Therefore, we control for these industry-specific effects by inserting dummy variables for industry groups in our models.

RESULTS

Descriptive results

Panel A of Table 5 presents the percentage of disclosing companies for each reporting item by superior and poor sustainability performers and for the full sample. For the full sample, minimum and maximum values are defined by R-S1 and R-S7. The highest proportion of quantitative disclosure (88%), the lowest proportion of qualitative disclosure (12%) and non-disclosure (0%) refer to R-S1, whereas the lowest proportion of quantitative disclosure (7%) and the highest proportion of qualitative disclosure (72%) refer to R-S7. In addition, the highest proportion of non-disclosure is related to R-E6. Across all indicators, approximately 50% of the companies on average opt for quantitative disclosure, whereas 38% (12%) opt for qualitative (non-)disclosure. Moreover, social reporting items are reported quantitatively more frequently than environmental reporting items. A similar distribution is obtained for the subsamples of superior and poor sustainability performers. For most reporting items, the proportion of quantitatively (qualitatively) disclosing firms is higher (lower) for superior sustainability performers than for poor sustainability performers, which is consistent with our expectations. However, both superior and poor sustainability performers report more quantitatively than qualitatively on average across all indicators.

Panel B of Table 5 reports the descriptive statistics on *CSDquan* and *CSDqual* by industry group. On average, quantitative (higher quality) disclosure is lowest for the ‘retail’ industry and highest for the ‘electricity’ industry, whereas qualitative (lower quality) disclosure is lowest for ‘banking and financial services’ and highest for the ‘retail’ industry. There is one firm that exclusively provides quantitative reporting (a car manufacturer) and one firm that does not disclose any reporting item quantitatively (a media company), but no firm discloses each reporting item qualitatively. On average, most industries use quantitative reporting to a greater extent than qualitative reporting, with the exception of ‘building materials, paper, steel’, ‘mining’, ‘retail’, ‘commercial services’, ‘aerospace and defense’, and ‘engineering and construction’. In general, reporting behavior is heterogeneous across industry groups, which supports the need to control for industry-specific effects.

Table 5: Descriptive statistics for *CSDquan* and *CSDqual* by indicator and by industry group

Panel A: Percentage of quantitatively/qualitatively disclosing companies per indicator

		<i>CSDquan</i>			<i>CSDqual</i>			<i>non-disclosure</i>		
		total	sup	poor	total	sup	poor	total	sup	poor
	observations	195	98	97	195	98	97	195	98	97
R-E1	materials used	40%	43%	37%	42%	39%	45%	18%	18%	18%
R-E2	energy consumption and renewables	39%	41%	38%	60%	58%	62%	1%	1%	0%
E-E3	water withdrawal	64%	68%	59%	30%	26%	34%	7%	6%	7%
R-E4	greenhouse gas emissions	55%	61%	49%	44%	38%	49%	1%	1%	1%
R-E5	ozone-depleting substances and other air emissions	42%	43%	41%	27%	23%	31%	31%	34%	28%
R-E6	water discharge	11%	9%	12%	47%	47%	47%	42%	44%	40%
R-E7	waste	49%	54%	44%	46%	40%	53%	5%	6%	3%
	mean environmental	43%	46%	40%	42%	39%	46%	15%	16%	14%
R-S1	workforce	88%	87%	89%	12%	13%	11%	0%	0%	0%
R-S2	employee turnover	67%	65%	69%	21%	21%	20%	12%	13%	11%
R-S3	collective bargaining agreements	50%	51%	48%	34%	35%	33%	16%	14%	19%
R-S4	safety and health	76%	83%	70%	21%	16%	26%	3%	1%	4%
R-S5	training	77%	80%	75%	22%	20%	24%	1%	0%	1%
R-S6	discrimination	32%	37%	28%	61%	57%	64%	7%	6%	8%
R-S7	child, forced, and compulsory labor	7%	8%	6%	72%	73%	71%	21%	18%	23%
	mean social	57%	59%	55%	35%	34%	35%	8%	8%	9%
	mean total	50%	52%	48%	38%	36%	41%	12%	12%	12%

Panel B: Descriptive statistics for CSDquan and CSDqual by industry group

		<i>CSDquan</i>					<i>CSDqual</i>			
		n	mean	sd	min	max	mean	sd	min	max
1	chemicals	10	7.90	1.85	5.00	11.00	5.50	1.78	3.00	9.00
2	building materials, paper, steel	16	6.38	3.28	1.00	12.00	6.56	2.45	2.00	10.00
3	mining	11	6.36	2.62	1.00	10.00	7.18	2.23	4.00	12.00
4	advertising, entertainment, media	11	5.73	3.23	0.00	11.00	5.82	2.36	2.00	10.00
5	telecommunication	11	8.18	2.68	4.00	12.00	3.55	1.37	2.00	6.00
6	computers and software	5	6.40	2.51	4.00	10.00	5.20	1.10	4.00	6.00
7	car manufacturers	8	8.13	3.44	3.00	14.00	5.38	2.92	0.00	10.00
8	retail	5	4.00	1.41	2.00	5.00	8.40	1.14	7.00	10.00
9	food and beverages	9	6.33	2.50	3.00	11.00	6.11	2.42	2.00	10.00
10	commercial services	7	5.00	2.89	3.00	11.00	6.71	2.21	3.00	10.00
11	pharmaceuticals	8	8.13	2.47	5.00	11.00	5.13	1.89	3.00	7.00
12	oil and gas	8	7.63	2.67	4.00	11.00	5.63	2.33	3.00	9.00
13	banking and financial services	30	8.67	2.52	3.00	13.00	3.43	1.52	1.00	7.00
14	insurance	10	7.40	2.67	4.00	11.00	4.60	1.96	2.00	9.00
15	REITS	10	5.30	1.89	2.00	8.00	4.60	0.70	4.00	6.00
16	aerospace and defense	8	5.13	1.55	3.00	7.00	6.63	2.13	4.00	11.00
17	engineering and construction	11	4.55	2.38	2.00	10.00	7.45	1.69	4.00	10.00
18	electricity	7	9.71	2.06	6.00	12.00	3.71	1.50	2.00	6.00
19	gas, water, and electricity distribution	10	7.70	2.71	4.00	11.00	5.30	2.31	2.00	9.00

For each reporting item, Panel A presents the percentage of superior and poor sustainability performers that report quantitatively or qualitatively or that do not report. Superior sustainability performers (sup) are firms with *CSP* equal to or above the sample median. Poor sustainability performers (poor) are firms with *CSP* below the sample median. Panel B presents descriptive statistics for *CSDquan* (*CSDqual*) by industry group.

Panel A of Table 6 presents descriptive statistics for the variables used in the regression analysis. The mean value for *CSDquan* is higher than that of *CSDqual*, which indicates that on average sample firms slightly prefer quantitative (higher quality) to qualitative (lower quality) disclosure, as defined in Table 1 and in the “CSD measurement scheme” section. *CSP* varies between 0 and 1, with a mean value slightly above 0.5. Approximately 68% of the sample firms adopt a sustainability assurance statement, which is a rather high proportion compared with previous studies. For instance, using a panel of Fortune Global 250 firms for 2008, Perego and Kolk (2012) report that 56% of sustainability reports include assurance statements. Because of fundamental differences between the litigation traditions of the European and U.S. markets, the deviation might stem from differences in sample composition, which is restricted to European companies in our study. The deviation may also be attributable to different observation periods because external assurance has become more commonly employed by large companies in recent years (KPMG, 2011). The mean of 2.49 for *ORG* corresponds to a percentage of 81% (untabulated) of the sample firms that run group-wide organizational structures to coordinate their corporate sustainability activities. This proportion exceeds the observations in previous studies (e.g. Al-Tuwaijri et al., 2004, report 62% and Ruhnke and Gabriel, 2013, report 59%). However, these deviations likely originate from differing scales and definitions. Because *SIZE* is highly positively skewed, we employ log-transformations to the original data and report the log-transformed data. On average, our sample firms employ 66,037 employees and are thus relatively large companies. Financial performance – as measured by free cash flow on the basis of euro per employee – measures approximately 0.03 on average, and the high standard deviation of financial leverage reflects the diversity of the sample with respect to different industry groups.

Panel B of Table 6 presents the Pearson correlations between all variables in our models. As expected, there is a negative correlation between *CSDquan* and *CSDqual*, which indicates a close but not perfect relationship between these variables and the use of the non-disclosure option by sample firms. The correlation coefficient between *CSP* and *CSDquan* is positive (but not significant), whereas the correlation coefficient between *CSP* and *CSDqual* is negative (and significant). The predicted signs of our control variables are generally in accordance with the correlation statistics, except for *TOBIN*. As expected, there is a positive (negative) relationship between *ASSURE* and *CSDquan* (*CSDqual*), which indicates that companies with assured *CSD* disclose higher quality information. *ORG* is positively correlated with *CSDquan*, but the strength of this correlation is weak and there is no correlation between *ORG* and *CSDqual*. Moreover, *ORG* is positively (and significantly) correlated with both *ASSURE* and *SIZE*. The correlation coefficients of both *SIZE* and *FCF* are insignificant, which reflects the opposing theoretical considerations regarding these control variables. *LEV* is positively (negatively) and significantly correlated with *CSDquan* (*CSDqual*). Counterintuitively, *TOBIN* is negatively (positively) correlated with *CSDquan* (*CSDqual*) and there is a negative correlation between *ASSURE* and *TOBIN*, which indicates that a lower degree of information asymmetry accompanies more quantitative *CSD* and external assurance, respectively.

Table 6: Descriptive and correlation statistics for regression variables

<i>Panel A: Descriptive statistics</i>							
	n	mean	median	sd	25th percentile	75th percentile	
(1) <i>CSDquan</i>	195	6.98	7.00	2.88	5.00	9.00	
(2) <i>CSDqual</i>	195	5.38	5.00	2.32	4.00	7.00	
(3) <i>CSP</i>	195	0.55	0.57	0.19	0.42	0.70	
(4) <i>ASSURE</i>	195	0.68	1.00	0.47	0.00	1.00	
(5) <i>ORG</i>	195	2.49	2.00	1.10	2.00	4.00	
(6) <i>SIZE</i>	195	4.42	4.61	0.76	4.01	4.97	
(7) <i>FCF</i>	195	0.03	0.02	0.32	0.00	0.06	
(8) <i>LEV</i>	195	6.91	3.40	11.52	2.27	6.82	
(9) <i>TOBIN</i>	195	1.25	1.06	0.48	0.98	1.30	

<i>Panel B: Correlation statistics</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	1.0000							
(2)	−0.8042 (0.0000)	1.0000						
(3)	0.0974 (0.1757)	−0.1187 (0.0983)	1.0000					
(4)	0.3991 (0.0000)	−0.2165 (0.0004)	−0.1004 (0.1624)	1.0000				
(5)	0.1319 (0.0660)	0.0202 (0.7787)	−0.0143 (0.8423)	0.3151 (0.0000)	1.0000			
(6)	0.0753 (0.2191)	0.1326 (0.0300)	0.0612 (0.3956)	0.2317 (0.0001)	0.2370 (0.0009)	1.0000		
(7)	0.0525 (0.3923)	−0.0458 (0.4556)	0.1061 (0.1398)	−0.0424 (0.4894)	−0.0575 (0.4248)	0.2150 (0.0004)	1.0000	
(8)	0.1439 (0.0184)	−0.2227 (0.0002)	0.0014 (0.9850)	−0.0218 (0.7219)	−0.1345 (0.0608)	0.0109 (0.8596)	0.1608 (0.0083)	1.0000
(9)	−0.1394 (0.0225)	0.1349 (0.0273)	0.0779 (0.2791)	−0.1263 (0.0387)	0.0440 (0.5414)	−0.0695 (0.2570)	0.0481 (0.4332)	−0.2492 (0.0000)

(1) *CSDquan*, (2) *CSDqual*, (3) *CSP*, (4) *ASSURE*, (5) *ORG*, (6) *SIZE*, (7) *FCF*, (8) *LEV*, (9) *TOBIN*

Table 6 presents descriptive statistics (Panel A) and correlation statistics (Panel B) for the variables used in the regression analysis. Statistics are presented for the full sample of 195 firms. Panel B reports bivariate Pearson correlation coefficients and p-values (in parentheses) for a two-tailed test of statistical significance.

Results of the regression analyses

The results of multivariate regression analyses with robust standard errors (White, 1980) are presented in Table 7. The first set of columns corresponds to hypothesis H1 (*CSDquan*), and the second set of columns corresponds to hypothesis H2 (*CSDqual*). For each hypothesis we present three models. Model (1) contains only our main variable of interest *CSP* along with *ASSURE* and *ORG*, which capture the strategic orientation toward sustainability issues. In model (2) we include all control variables except country and industry group dummies, whereas model (3) corresponds to our full model.

Table 7: Regression results

$$CSDquan = \beta_0 + \beta_1 CSP + \beta_2 ASSURE + \beta_3 ORG + \beta_4 SIZE + \beta_5 FCF + \beta_6 LEV + \beta_7 TOBIN \\ + \sum_{i=8}^{14} \beta_i COUNTRY + \sum_{j=15}^{33} \beta_j INDUSTRY + \varepsilon$$

$$CSDqual = \beta_0 + \beta_1 CSP + \beta_2 ASSURE + \beta_3 ORG + \beta_4 SIZE + \beta_5 FCF + \beta_6 LEV + \beta_7 TOBIN \\ + \sum_{i=8}^{14} \beta_i COUNTRY + \sum_{j=15}^{33} \beta_j INDUSTRY + \varepsilon$$

	<i>CSDquan</i> models			<i>CSDqual</i> models		
	(1)	(2)	(3)	(1)	(2)	(3)
<i>Intercept</i>	3.6605*** (4.6474)	3.7827*** (2.6171)	3.5434* (1.8771)	6.7406*** (9.8235)	5.2248*** (4.3195)	5.0303*** (3.2326)
<i>CSP (+/-)</i>	2.6614** (2.5332)	2.7335** (2.5697)	3.3036*** (3.3144)	-1.9094** (-2.0862)	-2.0598** (-2.3139)	-2.6497*** (-3.2248)
<i>ASSURE (+/-)</i>	2.5350*** (5.7187)	2.6020*** (5.7167)	2.1922*** (4.9277)	-1.2620*** (-3.2680)	-1.4574*** (-3.8263)	-1.1729*** (-3.1983)
<i>ORG (+/-)</i>	0.0533 (0.2859)	0.1401 (0.7311)	0.2045 (1.0171)	0.2051 (1.2627)	0.0593 (0.3697)	-0.0388 (-0.2340)
<i>SIZE</i>		-0.1397 (-0.4964)	-0.1051 (-0.2986)		0.5887** (2.4989)	0.4521 (1.5580)
<i>FCF</i>		0.5415 (0.8336)	-0.4773 (-0.6984)		-0.2333 (-0.4291)	0.1162 (0.2062)
<i>LEV (+/-)</i>		0.0409** (2.2799)	0.0141 (0.6805)		-0.0622*** (-4.1477)	-0.0199 (-1.1681)
<i>TOBIN (+/-)</i>		-0.0737 (-0.1739)	0.1696 (0.3690)		-0.0396 (-0.1117)	-0.3283 (-0.8667)
<i>Country dummies</i>	NO	NO	YES	NO	NO	YES
<i>Industry group dummies</i>	NO	NO	YES	NO	NO	YES
Observations	195	195	195	195	195	195
Adjusted R²	0.1648	0.1770	0.3635	0.0530	0.1407	0.3349
F-Statistic	13.7628***	6.96***	4.5747***	4.62***	5.5382***	4.1511***

Table 7 reports ordinary least squares coefficient estimates and t-statistics (in parentheses) based on Huber-White robust standard errors. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels (two-tailed), respectively.

With respect to hypothesis H1, the results of the multivariate regression analysis suggest a positive and significant association between *CSP* and *CSDquan* in all three models. This result is consistent with the predictions derived from VDT and indicates that superior sustainability performers disclose sustainability information quantitatively (higher quality) rather than qualitatively (lower quality) because this type of information is more reliable and comparable. By disclosing primarily quantitative (higher quality) information, these companies actively reveal their superior performance type to the market and are therefore able to distinguish themselves from poor sustainability performers. With respect to our control variables, only *ASSURE* is significant in all models, and *LEV* is significant in the reduced model. In addition to sustainability performance, a company's strategic orientation toward sustainability topics, which is captured by *ASSURE* and *ORG*, also has an impact on a company's disclosure strategy. Controlling for all other factors, companies with an active strategic posture more frequently opt for quantitative (higher quality) than qualitative (lower quality) disclosure. The insig-

nificant results for *SIZE* and *FCF* may reflect the opposing theoretical considerations on the relationship between CSD and these control variables. With respect to *SIZE*, the insignificant finding might also be attributed to our sample, which consists of relatively large firms. The insignificant findings for *FCF* are prevalent in the literature (Clarkson et al., 2011; Dawkins & Fraas, 2011). Additionally, the positive relationship for *LEV* in the reduced model indicates that companies with higher leverage are more likely to disclose high quality CSD. Financial leverage is largely industry-specific and the change in significance for *LEV* from model (2) to model (3) may stem from the inclusion of industry group dummies. Finally, *TOBIN* is insignificant, which is consistent with the prior literature (Clarkson et al., 2008; Clarkson et al., 2011).

The findings for hypothesis H2 reveal a negative and significant relationship between *CSP* and *CSDqual* (lower quality) in all three models. The reasoning of LT is supported, which indicates that poor sustainability performers disclose sustainability information qualitatively (lower quality) rather than quantitatively (higher quality) to manipulate public perceptions regarding their CSP. Because lower quality information typically lacks reliability and comparability, it is particularly useful for disguising a company's poor sustainability performance while still contributing to a sustainable company image. With respect to the control variables, *ASSURE* is significant in all models, whereas *SIZE* and *LEV* are significant in the reduced model. The negative result for *ASSURE* is evidence that a passive strategic orientation toward sustainability issues results in predominantly qualitative disclosures. Similar to our multivariate results on *CSDquan*, this relationship holds for the existence of external control mechanisms (*ASSURE*) and is insignificant for internal organizational structures (*ORG*). We observe a positive relationship between *SIZE* and *CSDqual* in the reduced model, which indicates that an increase in firm size might be associated with more qualitative disclosures, which is consistent with the reasoning of LT. In accordance with the prior literature, the relationship between *FCF* and *CSDqual* is insignificant. Consistent with our expectations, we observe a negative coefficient for *LEV* in the reduced model, which indicates that creditors may be less willing to accept qualitative (lower quality) information and may require higher levels of transparency with an increase in financial leverage. Finally, the relationship between *TOBIN* and *CSDqual* remains insignificant in all models.

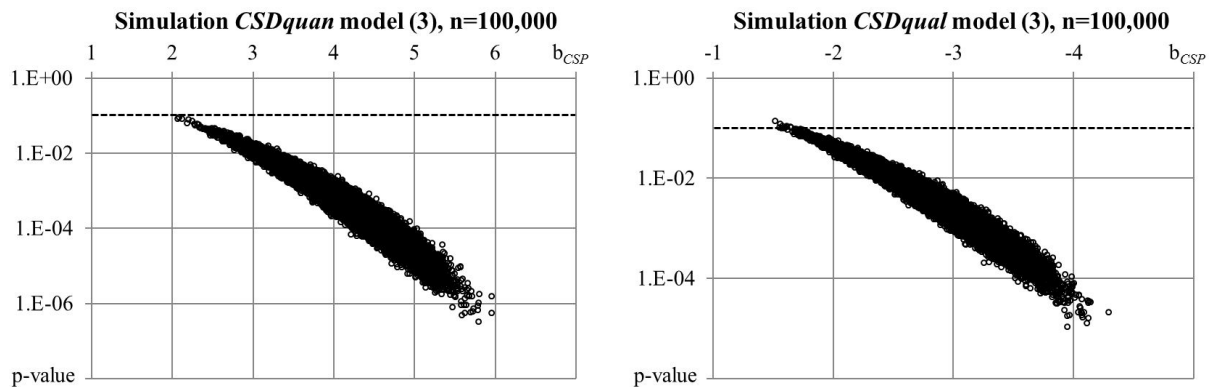
Supplemental analyses

We perform several variations to assess the robustness of our findings. First, we rerun the models using alternative proxies for firm size and financial performance. We use a log-transformation of market capitalization as an alternative proxy for firm size and both return on assets (ROA) and return on equity (ROE) for financial performance. The F-statistics and significance of the regression models – in addition to the direction and significance of the independent variables – are similar to those of the initial model, which supports the stability of our model (untabulated results). Second, we divide our sample into poor and superior sustainability performers based on median *CSP*. In accordance with our hypotheses, the mean value of *CSDquan* (*CSDqual*) is significantly higher (lower) for superior sustainability performers (t-tests of means, significance levels one-tailed, results untabulated). Third, we check the ro-

business of our results with respect to the winsorization of our original CSP data. In our baseline models, CSP data are winsorized at the 10% level. Although the winsorization procedure limits the influence of spurious outliers, it also reduces the variance in our raw data and its originality. Therefore, we rerun our models with CSP data winsorized at the 5% level and at the 1% level, and our results are similar to those of the initial model. In particular, both the direction and significance of the *CSP* variable remain unchanged. Thus, the results obtained are not significantly influenced by our handling of spurious outliers (results untabulated).

Finally, we account for the problem of missing data for certain performance indicators in the sample. According to our measurement scheme for *CSP*, these missing values are replaced with the mean values of the remaining indicators for each dimension. Following LT, it can be assumed that this approach might weaken the results of our study because poor performance values, in particular, are generally unreported. However, theoretical considerations might be misleading, and some firms may also have unobserved incentives to withhold the disclosure of superior performance data. To assess the robustness of our results with respect to missing values, we cover a wide range of possible scenarios using the Monte Carlo simulation method (Metropolis & Ulam, 1949). Beginning with winsorized original data, missing values are replaced by simulated values.¹¹ The simulated values are randomly drawn from a uniform distribution and separately rescaled for each industry group within an interval of $[0.8 \times \text{minimum per industry}; 1.2 \times \text{maximum per industry}]$. The additional range of 20% accounts for the possibility that missing data constitute unobserved extremes in our sample. The subsequent data rescaling and aggregation procedure is identical to our original method. The new dataset containing observed and simulated values is used to run our main regression models to test hypothesis 1 (hypothesis 2). The procedure of random sampling, data aggregation and hypothesis testing is repeated 100,000 times, and the results are presented in Figure 1.

Figure 1: Results from the Monte Carlo simulation for the *CSP* coefficients



Each mark represents a simulated regression coefficient and the corresponding p-value for *CSDquan* (left side) and *CSDqual* (right side). For all simulated scenarios, the signs of the regression coefficients for *CSP* remain positive (negative). With respect to *CSDquan*, 98.02% of all scenarios yield statistically significant results at the 1% level (1.97% at the 5% level and 0.01% at the 10% level). There is no scenario with insignificant results, which is indicated by

¹¹ Specifically, 96.21% (254/264) of the missing values are simulated. Ten gaps in P-S2 remain in the insurance industry because no data are provided.

the broken line in Figure 1. For *CSDqual*, 82.18% of all scenarios are significant at the 1% level (17.54% at the 5% level and 0.28% at the 10% level). In this case, only 0.01% of the results from all scenarios are insignificant (marks above the broken line). Overall, the results of the Monte Carlo simulation comprehensively support the robustness of our findings with respect to missing performance values in our data.

CONCLUSIONS

Recent research points toward an overlap between VDT and LT in explaining the sustainability reporting behavior of firms. Although previous studies contain preliminary support for this reasoning, there is an absence of robust empirical evidence in the literature. Our paper fills this research gap and thereby extends the prior literature in various ways. First, by concentrating on CSD quality instead of CSD quantity we are able to show the simultaneous applicability of predictions derived from both VDT and LT. In accordance with the reasoning of VDT, we argue that superior sustainability performers disclose sustainability information quantitatively rather than qualitatively because quantitative information is more reliable and comparable. In addition, we build on the reasoning of LT and predict a negative relationship between CSP and qualitative (lower quality) CSD because poor sustainability performers prefer to make lower quality disclosures to protect their image as sustainable firms. Second, our study expands the scope of previous studies that have focused exclusively on the environmental dimension of sustainability by also accounting for the social dimension. In this manner, we also respond to the call of several researchers and present improved measurement approaches for both CSP and CSD quality. We distinguish between quantitative (higher quality) disclosure, qualitative (lower quality) disclosure and non-disclosure for seven environmental and seven social disclosure items and aggregate the values into a total score for quantitative (higher quality), qualitative (lower quality) and non-disclosure. CSP is measured by four corresponding environmental and four corresponding social performance indicators. We refer directly to the data provided by our sample companies and aggregate these data into an overall performance score on an industry group basis. Third, the results from multivariate analyses of hand-collected data from 195 publicly traded European companies corroborate our theoretical reasoning. We observe a positive relationship between CSP and quantitative (higher quality) CSD and a negative relationship between CSP and qualitative (lower quality) CSD. The robustness of these findings is supported by several supplemental analyses. Additionally, our analysis provides insights into other determinants of CSD quality, which indicate that the distinction between quantitative (higher quality) and qualitative (lower quality) CSD prevails, in particular, with respect to an active strategic orientation toward sustainability and financial leverage.

In addition to the contributions to the academic literature, our study also has practical implications. First, the results reveal that there is currently not only a problem of non-reporting – as has been noted in criticism from institutional investors – but also a problem of different reporting qualities. To ensure the comparability of reported data, a precise and binding regulatory framework for the contents of sustainability reports is necessary. The development of such a framework would facilitate the external assessment of a company's sustainability perfor-

mance and its comparison with other peer companies. Second, in contrast to previous studies, findings from our study emphasize the importance of not only the environmental but also the social dimension of sustainability for the general public and the capital markets. Companies must be aware of the increasing importance of social disclosure and must also be prepared to provide relevant and high quality information.

The results of this paper are subject to certain limitations. First, the generalizability of our findings depends on the sample and time period of our study. To provide recent empirical results, we concentrate on the 2011 reporting period. Potential bias may arise from the sample composition of our study, which predominantly consists of large and publicly listed companies. Therefore, our results may not hold for small firms and/or firms less oriented to the capital markets. A second caveat of our study is linked to the problem of non-reporting companies. Our CSP measurement scheme requires that a company reveals performance data; as a consequence, our results cannot account for companies excluded from the sample because of insufficient data. Biases may also arise from companies that were included and partly withheld data. Nevertheless, results from a Monte Carlo simulation of these missing performance values do not indicate reasonable concerns with respect to the robustness of our initial findings.

Finally, our paper suggests directions for future empirical research. Our measures of CSD quality and CSP can be employed to analyze various questions that remain open in the field of sustainability disclosure. For instance, the results of our study provide preliminary evidence regarding the relevance of high quality sustainability disclosure for capital market participants. A different research design is necessary to test whether high quality CSD is indeed appraised by capital market participants and whether it affects firm value.¹² The results from previous investigations on the value relevance of CSD in general (Clarkson et al., 2013; Dhaliwal, Radhakrishnan, Tsang, & Yang, 2012) may be confounded by the overlapping effects of different CSD qualities. Thus, the application of our measurement schemes for CSD quality and CSP may add a new perspective to the ongoing discussion in this field of research.

¹² This research question must be separated from an overwhelming number of investigations on the relationship between CSP and financial performance (for an overview, see Dixon-Fowler, Slater, Johnson, Ellstrand, and Romi, 2013; Orlitzky, Schmidt, and Rynes, 2003).

Appendix: Previous empirical studies on the relationship between CSD and CSP

Author (year)	n country	CSD measurement	CSP measurement
<i>Cho et al.</i> (2012)	92 US	<ul style="list-style-type: none"> analysis of sustainability reports and annual reports based on <i>Clarkson et al.</i> (2008) 	<ul style="list-style-type: none"> environmental impact score by Newsweek/Trucost: includes air emissions, water use and waste
<i>Clarkson et al.</i> (2011)	51 AU	<ul style="list-style-type: none"> analysis of sustainability reports and annual reports based on <i>Clarkson et al.</i> (2008) 	<ul style="list-style-type: none"> toxic releases per sale (90 risk-weighted substances)
<i>Dawkins & Fraas</i> (2011)	344 US	<ul style="list-style-type: none"> reaction to a questionnaire by the Carbon Disclosure Project three groups: complete answer, incomplete response/ provided alternative information, no response/refused to respond 	<ul style="list-style-type: none"> rating by KLD: strengths and concerns rating of carbon dioxide emissions
<i>de Villiers & van Staden</i> (2011)	120 US	<ul style="list-style-type: none"> analysis of annual reports and corporate websites number of sentences addressing the environment and not related to litigation 	<ul style="list-style-type: none"> rating by KLD: strengths score minus concerns score
<i>Aerts & Cormier</i> (2009)	158 CA/US	<ul style="list-style-type: none"> analysis of annual reports 39 items in six categories: expenditures/risks, compliance, pollution abatement, sustainable development, remediation/contamination, environmental management weighting based on <i>Wiseman</i> (1982) 	<ul style="list-style-type: none"> toxic releases per sale
<i>Clarkson et al.</i> (2008)	191 US	<ul style="list-style-type: none"> analysis of sustainability reports 95 items in seven categories related to GRI guidelines (16 soft/easy to mimic, 79 hard/difficult to mimic) 0/1-classification 	<ul style="list-style-type: none"> toxic waste recycling rate (rank) toxic waste per sale (rank)
<i>Cho & Patten</i> (2007)	100 US	<ul style="list-style-type: none"> analysis of annual reports based on <i>Patten</i> (2002) separate analysis of four monetary and four non-monetary elements 	<ul style="list-style-type: none"> rating by KLD: “worse” or “better” performers depending on identified environmental concerns
<i>Cho et al.</i> (2006)	119 US	<ul style="list-style-type: none"> analysis of annual reports based on <i>Patten</i> (2002) 	<ul style="list-style-type: none"> rating by KLD: total score environmental concerns
<i>de Villiers & van Staden</i> (2006)	≥137 ZA	<ul style="list-style-type: none"> analysis of annual reports based on <i>Wiseman</i> (1982) general vs. specific disclosures 0/1-classification separate analysis of eight general disclosure items regarding employees 	<ul style="list-style-type: none"> not explicitly measured relationship is assumed
<i>Al-Tuwaijri et al.</i> (2004)	198 US	<ul style="list-style-type: none"> analysis of annual reports four environmental indicators: toxic waste, financial penalties, cleanup responsibility for hazardous-waste sites, oil and chemical spills weighting: quantitative (3), non-quantitative but related to indicators (2), general qualitative (1), no disclosure (0) adjusted for firm-specific relevance 	<ul style="list-style-type: none"> toxic waste recycling rate toxic waste per sale firm’s attitude toward environmental concerns (3 indicators)
<i>Patten</i> (2002)	131 US	<ul style="list-style-type: none"> analysis of annual reports four monetary and four non-monetary elements 0/1-classification additionally: line count 	<ul style="list-style-type: none"> toxic releases per sale
<i>Hughes et al.</i> (2001)	51 US	<ul style="list-style-type: none"> analysis of annual reports based on <i>Wiseman</i> (1982) five additional elements weighting: quantitative (4), descriptive (3), vague (2), immaterial (1), no disclosure (0) 	<ul style="list-style-type: none"> three groups (good, mixed, poor) based on ranking by Council on Economic Priorities: emissions to air and water, prevention programs
<i>Bewley & Li</i> (2000)	188 CA	<ul style="list-style-type: none"> analysis of annual reports based on <i>Wiseman</i> (1982) additionally: financial vs. non-financial information 	<ul style="list-style-type: none"> pollution propensity (high/low) depending on industry toxic releases reported to federal database (yes/no)

Author (year)	n country	CSD measurement	CSP measurement
<i>Fekrat et al.</i> (1996)	26 US	<ul style="list-style-type: none"> • analysis of annual reports • based on <i>Wiseman</i> (1982) 	<ul style="list-style-type: none"> • based on <i>Ingram & Frazier</i> (1980)
<i>Rockness</i> (1985)	26 US	<ul style="list-style-type: none"> • field experiment • participants' perception of statements in annual reports 	<ul style="list-style-type: none"> • based on <i>Ingram & Frazier</i> (1980)
<i>Wiseman</i> (1982)	26 US	<ul style="list-style-type: none"> • analysis of annual reports • four categories containing eighteen items: economic, litigation, pollution, other information • weighting: quantitative (3), non-quantitative but related (2), general (1), no disclosure (0) • additionally: line count 	<ul style="list-style-type: none"> • based on <i>Ingram & Frazier</i> (1980)
<i>Ingram & Frazier</i> (1980)	40 US	<ul style="list-style-type: none"> • analysis of annual reports • four dimensions broken down into 20 categories: evidence (e.g., monetary, non-monetary, qualitative), time (past, present, future), specificity (specific, general), theme (e.g., public interest, economic consequences) • check sentence by sentence • results per category and dimension (total score) 	<ul style="list-style-type: none"> • index by Council on Economic Priorities: emissions to air and water, prevention programs

(AU) Australia, (CA) Canada, (US) United States, (ZA) South Africa, (n) sample size

REFERENCES

- Abdel-Khalik, A. R. (1993). Why do private companies demand auditing? A case for organizational loss of control. *Journal of Accounting, Auditing, and Finance*, 8(1), 31–52.
- Adams, C. A. (2002). Internal organisational factors influencing corporate social and ethical reporting: Beyond current theorising. *Accounting, Auditing & Accountability Journal*, 15(2), 223–250.
- Aerts, W., & Cormier, D. (2009). Media legitimacy and corporate environmental communication. *Accounting, Organizations and Society*, 34(1), 1–27.
- Aerts, W., Cormier, D., & Magnan, M. (2008). Corporate environmental disclosure, financial markets and the media: An international perspective. *Ecological Economics*, 64(3), 643–659.
- Al-Tuwaijri, S. A., Christensen, T. E., & Hughes, K. E. (2004). The relations among environmental disclosure, environmental performance, and economic performance: A simultaneous equations approach. *Accounting, Organizations and Society*, 29(5–6), 447–471.
- Ballou, B., Casey, R. J., Grenier, J. H., & Heitger, D. L. (2012). Exploring the strategic integration of sustainability initiatives: Opportunities for accounting research. *Accounting Horizons*, 26(2), 265–288.
- Ballou, B., Heitger, D. L., & Landes, C. E. (2006). The future of corporate sustainability reporting. *Journal of Accountancy*, 202(6), 65–74.
- Bewley, K., & Li, Y. (2000). Disclosure of environmental information by Canadian manufacturing companies: A voluntary disclosure perspective. *Advances in Environmental Accounting & Management*, 1, 201–226.

- Branco, M. C., & Rodrigues, L. L. (2008). Factors influencing social responsibility disclosure by Portuguese companies. *Journal of Business Ethics*, 83(4), 685–701.
- Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy*, 18(1), 125–169.
- Chin, M. K., Hambrick, D. C., & Treviño, L. K. (2013). Political ideologies of CEOs: The influence of executives' values on corporate social responsibility. *Administrative Science Quarterly*, 58(2), 197–232.
- Chiu, S. C., & Sharfman, M. (2011). Legitimacy, visibility, and the antecedents of corporate social performance: An investigation of the instrumental perspective. *Journal of Management*, 37(6), 1558–1585.
- Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? An empirical investigation of corporate environmental reputation. *Accounting, Organizations and Society*, 37(1), 14–25.
- Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, Organizations and Society*, 32(7–8), 639–647.
- Cho, C. H., Patten, D. M., & Roberts, R. W. (2006). Corporate political strategy: An examination of the relation between political expenditures, environmental performance, and environmental disclosure. *Journal of Business Ethics*, 67(2), 139–154.
- Clarkson, P. M., Fang, X., Li, Y., & Richardson, G. D. (2013). The relevance of environmental disclosures: Are such disclosures incrementally informative? *Journal of Accounting and Public Policy*, 32(5), 410–431.
- Clarkson, P. M., Li, Y., & Richardson, G. D. (2004). The market valuation of environmental expenditures by pulp and paper companies. *The Accounting Review*, 79(2), 329–353.
- Clarkson, P. M., Li, Y., Richardson, G. D., & Vasvari, F. P. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, 33(4–5), 303–327.
- Clarkson, P. M., Overell, M. B., & Chapple, L. (2011). Environmental reporting and its relation to corporate environmental performance. *Abacus*, 47(1), 27–60.
- Cormier, D., & Magnan, M. (2003). Environmental reporting management: A continental European perspective. *Journal of Accounting and Public Policy*, 22(1), 43–62.
- Cormier, D., Magnan, M., & van Velthoven, B. (2005). Environmental disclosure quality in large German companies: Economic incentives, public pressures or institutional conditions? *European Accounting Review*, 14(1), 3–39.
- Cowen, S. S., Ferreri, L. B., & Parker, L. D. (1987). The impact of corporate characteristics on social responsibility disclosure: A typology and frequency-based analysis. *Accounting, Organizations and Society*, 12(2), 111–122.

- Davis, K. (1973). The case for and against business assumption of social responsibilities. *Academy of Management Journal*, 16(2), 312–322.
- Dawkins, C., & Fraas, J. W. (2011). Coming clean: The impact of environmental performance and visibility on corporate climate change disclosure. *Journal of Business Ethics*, 100(2), 303–322.
- de Villiers, C., & van Staden, C. J. (2006). Can less environmental disclosure have a legitimising effect? Evidence from Africa. *Accounting, Organizations and Society*, 31(8), 763–781.
- de Villiers, C., & van Staden, C. J. (2011). Where firms choose to disclose voluntary environmental information. *Journal of Accounting and Public Policy*, 30(6), 504–525.
- Deegan, C. (2002). Introduction: The legitimizing effect of social and environmental disclosures – A theoretical foundation. *Accounting, Auditing & Accountability Journal*, 15(3), 282–311.
- Dhaliwal, D. S., Radhakrishnan, S., Tsang, A., & Yang, Y. G. (2012). Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The Accounting Review*, 87(3), 723–759.
- Dixon-Fowler, H. R., Slater, D. J., Johnson, J. L., Ellstrand, A. E., & Romi, A. M. (2013). Beyond “does it pay to be green?” A meta-analysis of moderators of the CEP–CFP relationship. *Journal of Business Ethics*, 112(2), 353–366.
- Dowling, J., & Pfeffer, J. (1975). Organizational legitimacy: Social values and organizational behavior. *The Pacific Sociological Review*, 18(1), 122–136.
- Elkington, J. (1997). *Cannibals with forks: The triple bottom line of 21st century business*. Oxford: Capstone.
- FASB (2010). *Statement of financial accounting concepts no. 8: Conceptual framework for financial reporting*. Norwalk, CT: FASB.
- Fekrat, M. A., Inclan, C., & Petroni, D. (1996). Corporate environmental disclosures: Competitive disclosure hypothesis using 1991 annual report data. *The International Journal of Accounting*, 31(2), 175–195.
- Fifka, M. S. (2013). Corporate responsibility reporting and its determinants in comparative perspective – A review of the empirical literature and a meta-analysis. *Business Strategy and the Environment*, 22(1), 1–35.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston, MA: Cambridge University Press.
- Gray, R. (2006). Social, environmental and sustainability reporting and organisational value creation? Whose value? Whose creation? *Accounting, Auditing and Accountability Journal*, 19(6), 793–819.
- GRI (2011a). *Sustainability reporting guidelines version 3.1*. The Netherlands: GRI.
- GRI (2011b). *Indicator protocols set environment (EN)*. The Netherlands: GRI.

- GRI (2011c). *Indicator protocols set labor practices and decent work (LA)*. The Netherlands: GRI.
- GRI (2011d). *Indicator protocols set human rights (HR)*. The Netherlands: GRI.
- GRI (2013). *G4 reporting principles and standard disclosures*. The Netherlands: GRI.
- Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1–3), 405–440.
- Hong, Y., & Andersen, M. L. (2011). The relationship between corporate social responsibility and earnings management: An exploratory study. *Journal of Business Ethics*, 104(4), 461–471.
- Horváthová, E. (2012). The impact of environmental performance on firm performance: Short-term costs and long-term benefits? *Ecological Economics*, 84, 91–97.
- Hughes, S. B., Anderson, A., & Golden, S. (2001). Corporate environmental disclosures: Are they useful in determining environmental performance? *Journal of Accounting and Public Policy*, 20(3), 217–240.
- Ingram, R. W., & Frazier, K. B. (1980). Environmental performance and corporate disclosure. *Journal of Accounting Research*, 18(2), 614–622.
- Knechel, W. R., Salterio, S. E., & Ballou, B. (2007). *Auditing: Assurance and risk* (3rd ed.). Mason, OH: Thomson/South-Western.
- Kolk, A. (2003). Trends in sustainability reporting by the Fortune Global 250. *Business Strategy and the Environment*, 12(5), 279–291.
- Kolk, A., Walhain, S., & van de Weteringen, S. (2001). Environmental reporting by the Fortune Global 250: Exploring the influence of nationality and sector. *Business Strategy and the Environment*, 10(1), 15–28.
- KPMG (2011). *KPMG international survey of corporate responsibility reporting 2011*. Switzerland: KPMG.
- Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Los Angeles, CA: Sage.
- Li, Y., Richardson, G. D., & Thornton, D. B. (1997). Corporate disclosure of environmental liability information: Theory and evidence. *Contemporary Accounting Research*, 14(3), 435–474.
- McWilliams, A., & Siegel, D. (2001). Corporate social responsibility: A theory of the firm perspective. *Academy of Management Review*, 26(1), 117–127.
- Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, 44(247), 335–341.
- Neu, D., Warsame, H., & Pedwell, K. (1998). Managing public impressions: Environmental disclosures in annual reports. *Accounting, Organizations and Society*, 23(3), 265–282.

- O'Donovan, G. (2002). Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. *Accounting, Auditing & Accountability Journal*, 15(3), 344–371.
- Orij, R. (2010). Corporate social disclosures in the context of national cultures and stakeholder theory. *Accounting, Auditing & Accountability Journal*, 23(7), 868–889.
- Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies*, 24(3), 403–441.
- Patten, D. M. (1991). Exposure, legitimacy, and social disclosure. *Journal of Accounting and Public Policy*, 10(4), 297–308.
- Patten, D. M. (2002). The relation between environmental performance and environmental disclosure: A research note. *Accounting, Organizations and Society*, 27(8), 763–773.
- Perego, P., & Kolk, A. (2012). Multinationals' accountability on sustainability: The evolution of third-party assurance of sustainability reports. *Journal of Business Ethics*, 110(2), 173–190.
- Porter, M. E., & Kramer, M. R. (2011). Creating shared value. *Harvard Business Review*, 89(1–2), 62–77.
- Reuters (2011). Investors demand improved sustainability reporting. Retrieved 09/12/2014, from <http://www.reuters.com/article/2011/02/21/investors-sustainability-idUSLDE71K0W120110221>.
- Roberts, R. W. (1992). Determinants of corporate social responsibility disclosure: An application of stakeholder theory. *Accounting, Organizations and Society*, 17(6), 595–612.
- Rockness, J. W. (1985). An assessment of the relationship between US corporate environmental performance and disclosure. *Journal of Business Finance & Accounting*, 12(3), 339–354.
- Ruhnke, K., & Gabriel, A. (2013). Determinants of voluntary assurance on sustainability reports: An empirical analysis. *Journal of Business Economics*, 83(9), 1063–1091.
- Russo, M. V., & Fouts, P. A. (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of Management Journal*, 40(3), 534–559.
- Sethi, S. P. (1978). Advocacy advertising – The American experience. *California Management Review*, 21(1), 55–67.
- Stanny, E., & Ely, K. (2008). Corporate environmental disclosures about the effects of climate change. *Corporate Social Responsibility and Environmental Management*, 15(6), 338–348.
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610.
- Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, 33(1), 1–67.

- Ullmann, A. A. (1985). Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of U.S. firms. *Academy of Management Review*, 10(3), 540–557.
- van der Laan Smith, J., Adhikari, A., & Tondkar, R. H. (2005). Exploring differences in social disclosures internationally: A stakeholder perspective. *Journal of Accounting and Public Policy*, 24(2), 123–151.
- Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics*, 5(1), 179–194.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4), 817–838.
- Wiseman, J. (1982). An evaluation of environmental disclosures made in corporate annual reports. *Accounting, Organizations and Society*, 7(1), 53–63.
- Wood, D. J. (1991). Corporate social performance revisited. *Academy of Management Review*, 16(4), 691–718.

The Nature of Managerial Capture in Sustainability Assurance Engagements¹

ABSTRACT

This study investigates the nature of managerial capture in sustainability assurance engagements in an empirical setting. We hypothesize that managerial capture, i.e., a client's influence on the process and outcome of an assurance engagement, affects the assurance process and outcome in either a beneficial or a detrimental way with respect to transparency and accountability toward external stakeholders. To assess the nature of managerial capture, we concentrate on the relationship between the intensity of the sustainability assurance services requested by the client and a firm's corporate sustainability performance. Based on a sample of 122 European companies, our results reveal a negative relationship between assurance intensity and firms' sustainability performance, indicating the existence of a beneficial form of managerial capture.

INTRODUCTION

In the past two decades, global awareness of the importance of sustainability has rapidly increased. Stakeholders expect companies to act responsibly toward their natural and social environment and to demonstrate adherence to these expectations through corporate sustainability disclosure. Reporting on extra-financial activities, such as corporate sustainability and corporate social responsibility, "is now undeniably a mainstream global business practice" with 93 percent of the largest 250 companies worldwide publishing stand-alone or integrated sustainability reports in 2013 (KPMG, 2013, p. 22). Regarding the reasons for this development, previous research has shown that establishing legitimacy is a primary motive for a company's corporate sustainability disclosure, which affects both the quantity and the quality of the information presented (Cho, Guidry, Hageman, & Patten, 2012; Cho & Patten, 2007). Stakeholders are more likely to consider a company's actions to be legitimate, i.e., in accordance with society's expectations regarding sustainable business conduct (Suchman, 1995), if the information is disclosed in a convincing and credible way.

Primarily as a means to enhance the credibility of their sustainability reporting, firms have begun to seek external assurance of their sustainability disclosure and to publish the resulting assurance statements (Edgley, Jones, & Solomon, 2010; Park & Brorson, 2005; Simnett, Vanstraelen, & Chua, 2009). Indeed, KPMG reports that the proportion of assured reports among the largest 250 companies worldwide increased from 29 percent in 2002 to 59 percent in 2013 (KPMG, 2013). In addition, firms seek external assurance to improve their internal management and reporting systems and actual sustainability performance (Park & Brorson, 2005). The adoption of assurance, in conjunction with an assurance statement, can directly improve a firm's perceived legitimacy and indirectly lead to internal and performance-related improvements. Therefore, a sustainability assurance engagement can serve as a legitimization

¹ This paper is co-authored by Dr. Katrin Hummel, University of Zurich.

tool directed at stakeholders in addition to the sustainability disclosure itself (Darnall, Seol, & Sarkis, 2009; O'Dwyer, Owen, & Unerman, 2011).

Because of the generally voluntary nature of assurance, client companies largely determine the overall intensity of assurance engagements. Ultimately, the client pays for the assurance services; thus, the client can freely select those services that best serve her interests. Accordingly, concerns have been raised regarding whether assurance and assurance statements are merely another means of enhancing the public perception of a company's sustainability image, with detrimental effects on its accountability and transparency toward external stakeholders. The literature uses the term *managerial capture* (MC) to describe this critical picture of management's strategic influence on the process and outcome of assurance engagements (Ball, Owen, & Gray, 2000; Owen, Swift, Humphrey, & Bowerman, 2000). Previous empirical studies provide preliminary evidence for the existence of such a form of MC and generally derive their conclusions from the degree of stakeholder inclusivity in the assurance process and the quality of the assurance statement itself. For instance, O'Dwyer and Owen (2005, p. 224) conclude that "[...] stakeholder engagement with the reporting organization per se is often minimal which suggests to us that accountability to stakeholders is a low priority in these exercises. Hence, [...] there is still a large degree of management control over the reporting, and indeed assurance, process". Ball et al. (2000) draw similar conclusions. However, limited stakeholder inclusivity during the assurance engagement may partly stem from "the conferment of the assurator's mandate" (Manetti & Toccafondi, 2012, p. 373). In addition, recent findings from qualitative research reveal that assurers exhibit a high level of commitment to accountability (O'Dwyer et al., 2011). Especially major accountancy firms, which currently dominate two-thirds of the sustainability assurance market (KPMG, 2013), are likely to defend their commitment to accountability and to withstand the detrimental influence of their clients on the assurance process and its outcomes.

In light of the ambiguous findings in the literature and the recent developments in the market for assurance engagements, we propose a broader and more neutral understanding of MC as *a client's influence on the process and outcomes of assurance engagements*. We argue that the *nature* of MC can be *beneficial* or *detrimental* with respect to transparency and accountability toward external stakeholders. In the first case (*beneficial* MC), the client asks for more intensive assurance services to improve her internal sustainability-related systems and procedures and thus enhance its corporate sustainability performance (CSP), accountability, and transparency. In the opposite case (*detrimental* MC), the assurance engagement degenerates into a window-dressing activity that lacks serious attempts to improve the true state of a company's sustainability performance. In both scenarios, firms act under the assumption that an assurance statement increases the credibility of their sustainability disclosure, although the effect is likely to be stronger in the beneficial scenario. We introduce a company's CSP as a determining factor of the intensity of assurance services requested by the client in an assurance engagement. By investigating the relationship between CSP and assurance intensity (AI), we provide insights into the *nature* of MC. If the client captures the assurance services in a *detrimental* way, we expect to find a *positive* relationship between CSP and AI. This detrimental form of MC reflects a critical perspective on sustainability assurance because poor sustaina-

bility performers would obtain less intensive assurance to protect their external image without actually attempting to enhance their true state. Conversely, if a *beneficial* form of MC exists, we expect to find a *negative* relationship between CSP and AI. In this case, poor-performing companies seek more intensive assurance to improve their credibility, sustainability-related processes, and CSP.

Our investigation of the relationship between CSP and AI extends the literature in several ways. First, we introduce a broad understanding of MC and provide an example of the existence of a beneficial form of MC that advances recent qualitative evidence on the process involved in sustainability assurance engagements. Client firms indeed affect the intensity of assurance services; however, this influence can be beneficial for transparency and accountability depending on the firms' CSP. Second, to the best of our knowledge, our study is the first to investigate the nature of MC in an empirical setting. In so doing, we respond to the call from previous researchers for an empirical investigation of the "diffusion of assurance services in relation to a firm's social and environmental performance" (Kolk & Perego, 2010, p. 195). Moreover, we extend existing knowledge on the determinants of assurance adoption in general (Cho, Michelon, Patten, & Roberts, 2014; Kolk & Perego, 2010; Simnett et al., 2009) by providing insights into the determinants of AI for firms that adopt sustainability assurance. Our empirical results reveal a substantial impact of both firms' financial performance and firms' overall assurance and audit fees on AI. Third, our empirical model includes proxies for a provider's independence and competence, which allows us to support findings from the audit literature. Both determinants appear to be relevant not only for mandatory financial audits but also for voluntary sustainability assurance engagements.

The structure of the paper is as follows. In the next section, the related literature is reviewed, and hypotheses are developed. The subsequent section outlines our research design. In particular, the data sample and empirical model are presented and our main variables of interest, namely AI and CSP, are introduced. Then, descriptive statistics, findings from our main analyses, and robustness checks follow. Finally, we report our conclusions.

LITERATURE REVIEW AND HYPOTHESIS DEVELOPMENT

Prior evidence on managerial capture and theoretical background

Ongoing research on sustainability assurance includes three primary areas of interest.² Studies on the *adoption* of assurance predominantly focus on firm-, industry-, and country-specific determinants. Using an international sample of 2,113 companies, Simnett et al. (2009) report that companies in stakeholder-orientated countries (operationalized as code law countries) and companies with a greater need to enhance credibility (operationalized as affiliation to certain industries) are more likely to seek assurance for their sustainability reports. They also find that firm size positively influences the voluntary adoption of assurance. Similar results are obtained by Kolk and Perego (2010) and Cho et al. (2014), who report a significant impact of

² Moreover, experimental studies have examined report users' perceptions of the reliability of sustainability disclosure and assurance (Hodge, Subramaniam, & Stewart, 2009; Pflugrath, Roebuck, & Simnett, 2011).

country-level institutional factors in addition to industry affiliation on assurance adoption. Further, Darnall et al. (2009) show the role of internal and external stakeholders in the selection of various types of environmental audits. While these studies agree that enhanced credibility is a major reason for a company's voluntary adoption of assurance, Kolk and Perego (2010, p. 195) conclude that "future research should examine the quality of sustainability assurance statements rather than merely their adoption".

A second group of studies investigates the *outcomes* of sustainability assurance engagements in detail, i.e., assurance statement quality. The existence of MC is often highlighted, and researchers favor descriptive approaches based on content analysis to derive a critical picture of assurance engagements. Accordingly, Owen et al. (2000, p. 85) define MC as a "concept that sees management take control of the whole process [...] by strategically collecting and disseminating only the information it deems appropriate to advance the corporate image, rather than being truly transparent and accountable to the society it serves".³ Ball et al. (2000) analyze assurance statements from 53 European companies and ascertain a detrimental form of MC based on examinations of statement addressees, auditor independence, and the rigor of verification work. In another content analysis of 41 European statements, O'Dwyer and Owen (2005) similarly interpret the lack of transparency, accountability, and stakeholder inclusivity as evidence of the existence of a detrimental form of MC. Conclusions derived by Deegan, Cooper, and Shelly (2006) and Cooper and Owen (2007) complement these findings. However, Manetti and Toccafondi (2012, p. 373) note that there is "ample room for a variety of interpretations" and that MC should "be interpreted in the light of procedures for conferment of the assurors' mandate".

Valuable insights can therefore be gained at the third and most detailed level of investigation, at which the *process* of assurance engagements is examined. Studies in this area predominantly use qualitative methods such as interviews and case studies, and they find that sustainability assurance is an iterative process with close interactions between the client and its assurance provider (O'Dwyer, 2011; O'Dwyer et al., 2011; Park & Brorson, 2005). Although Edgley et al. (2010, p. 553) interpret the lack of stakeholder inclusivity as evidence of the existence of a detrimental form of MC, they conclude that "there is significant change underway in the area" and that "indirect stakeholder involvement [...] now represents good practice". This momentum is reflected in works by O'Dwyer (2011) and O'Dwyer et al. (2011), who reveal that assurance providers exhibit high commitment to accountability to external stakeholders. In particular, these authors emphasize the independence and power of providers in the assurance process. Some of these studies also examine companies' motives for sustainability assurance (Edgley et al., 2010; Park & Brorson, 2005; Simnett et al., 2009), identifying enhanced credibility, refined internal systems or processes, and improved CSP as primary reasons for adopting assurance. By contrast, companies may avoid adopting assurance because of the additional costs, the lack of perceived value added, and a lack of stakeholder demand (Park & Brorson, 2005).

³ (Power, 1991, 1996, 1997) also raises such concerns, noting the importance of substance in environmental assurance.

From a theoretical perspective, legitimacy theory (LT) and stakeholder theory (ST) contribute to the understanding of previous findings on the adoption, outcome, and process of sustainability assurance engagements (Darnall et al., 2009; O'Dwyer et al., 2011).⁴ Legitimacy is defined as “a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs, and definitions” (Suchman, 1995, p. 574). “If central stakeholders lose confidence in the firm’s performance, legitimacy may be withdrawn as the stakeholders refuse to provide their share of reciprocal benefits” and “if the firm cannot compensate for lost stakeholder benefits, it becomes ‘illegitimate’ and dies” (Wood, 1991, p. 697).⁵ Accordingly, firms that are perceived as non-sustainable by their stakeholders run the risks of harming their image, decreasing customer satisfaction, hampering their attraction to eligible employees, and facing litigation and regulatory constraints (Ameer & Othman, 2012; Ihlen & Roper, 2014; Kollman & Prakash, 2001; Wood, 1991). Firms may use a variety of strategies to cope with threatened legitimacy, such as shaping public perceptions through disclosure and altering business performance itself (Deegan, 2002; O'Donovan, 2002; Sethi, 1978). Sustainability assurance engagements can be regarded as legitimization tools that are in line with such strategies. The process of assurance helps firms to improve their internal systems, procedures, and, ultimately, CSP, which refers to the strategy of altering business performance. The signaling effect of assurance adoption in conjunction with assurance statements shape public perceptions by increasing a firm’s credibility and demonstrating the seriousness of its sustainability-related endeavors. In so doing, sustainability assurance engagements serve a firm’s self-interest of being perceived as legitimate and may simultaneously provide additional transparency and accountability to external stakeholders.

Overall, while our brief overview of the literature reveals that studies generally agree on the existence of MC and theoretical considerations provide a meaningful framework for MC, the determinants and thus the nature of MC remain unclear. Qualitative research has recently softened the initially critical picture of MC, revealing that negotiations regarding the terms of an assurance engagement and close interactions between the client and the assurance provider in the assurance process are common practice. Accordingly, we introduce a more general definition of MC as *a client’s influence on the process and outcome of assurance engagements*. Drawing on this broader and more neutral definition, we argue that MC can be beneficial or detrimental with respect to accountability and transparency toward external stakeholders.

Detrimental and beneficial managerial capture hypotheses

While previous studies primarily focus on the quality of assurance statements as an indicator of the existence of MC, we concentrate on the relationship between CSP and AI to assess whether a beneficial or detrimental form of MC exists. Because CSP improvements are a primary reason that firms adopt sustainability assurance, investigating the relationship between

⁴ LT refers to general societies in which companies exist, whereas ST refers to specific stakeholders within these societies “that can affect, or are affected by, the accomplishment of organizational purpose” (Freeman, 1984, p. 25).

⁵ Similar conclusions are also drawn by Davis (1973).

CSP and AI allows us to directly observe the nature of MC. We argue that a *positive* relationship between CSP and AI indicates a *detrimental* form of MC. Previous studies identify enhanced credibility as another major reason that firms seek external assurance (Edgley et al., 2010; Park & Brorson, 2005; Simnett et al., 2009). Because firms with poor CSP are particularly subject to legitimacy threats and because they experience greater difficulty in convincing their stakeholders of their sustainability, these companies may seek highly intensive assurance services to enhance their external credibility. However, if these poor-performing firms are not seriously interested in improving their poor internal systems, processes, and CSP, they will seek low-intensity assurance services. In this case, sustainability assurance merely serves as a window-dressing activity or merely reflects compliance with the common industry practice of adopting at least some type of assurance. The assurance engagement is regarded as a compulsory exercise that per se provides some credibility irrespective of its intensity. In this case, poor sustainability performers conceal the status of their internal systems and processes from their stakeholders by requesting lower intensity assurance services. Content-based analyses support this critical picture of assurance, indicating that management's control over the assurance process hampers the firm's accountability and transparency toward external stakeholders (Ball et al., 2000; Cooper & Owen, 2007; Deegan et al., 2006; O'Dwyer & Owen, 2005). In line with this argumentation, the following hypothesis is proposed:

Detrimental MC hypothesis: A positive relationship exists between corporate sustainability performance (CSP) and assurance intensity (AI).

However, if firms with poor CSP are interested in improving their internal systems, processes, and CSP, they can utilize intensive assurance services to support these goals with the further benefit of enhancing their credibility. Poor-performing companies seek more intensive assurance services that can also be delivered in an iterative and partly advisory way. In this case, a *beneficial* form of MC arises, which is reflected by a *negative* relationship between CSP and AI. This reasoning is also supported by recent developments in the assurance market, where major accountancy firms constituted two-thirds of the market in 2013 (KPMG, 2013). The increasing market shares of the major accountancy firms are accompanied by an increase in supply-side market power, which weakens clients' influence in the assurance engagement. Although assurance seekers – as purchasers of assurance services – still largely determine the intensity of engagements, the potential to capture the assurance process in a detrimental way is diminishing. Accordingly, recent qualitative research portrays assurance providers, particularly major accountancy firms, as independent, dedicated to the principle of stakeholder accountability, and highly risk averse with respect to the text of assurance statements (O'Dwyer, 2011; O'Dwyer et al., 2011). Against this background, we propose the following alternative hypothesis:

Beneficial MC hypothesis: A negative relationship exists between corporate sustainability performance (CSP) and assurance intensity (AI).

Our hypotheses are not mutually exclusive. Indeed, some firms may capture the assurance process in a detrimental way, whereas other firms may capture it in a beneficial way. In addition, firms may also opt for a mixed form of MC, as the motives for obtaining assurance ser-

vices and thus the two effects partially overlap and cancel one another out. Whether, on average, MC is beneficial or detrimental with respect to transparency and accountability toward external stakeholders remains an empirical question.

RESEARCH DESIGN

Assurance statement elements

While legally binding regulations on sustainability assurance are rare, assurance standards, namely, the International Standard on Assurance Engagements 3000 (ISAE 3000) and the AccountAbility 1000 Assurance Standard (AA1000AS), exist. Additionally, the Global Reporting Initiative (GRI) publishes recommendations on the process and outcome of sustainability assurance engagements. ISAE 3000 is provided by the International Auditing and Assurance Standards Board (IAASB), which is affiliated to the International Federation of Accountants (IFAC). The primary aim of ISAE 3000 is to “establish basic principles and essential procedures for [...] professional accountants [...] for the performance of assurance engagements other than audits or reviews of historical financial information” (IFAC, 2008, p. 4).⁶ The use of this standard is mandatory for assurance providers that are members of the IFAC, especially major accountancy firms, although the standard is not specifically designed for sustainability assurance engagements. Specific guidance on the assurance of sustainability-related disclosure is provided by the AA1000AS, which was developed by the international non-profit organization AccountAbility through a multi-stakeholder process. In addition to these standards, the GRI – originally known for its widely used guidelines on sustainability reporting – has begun to publish recommendations on the selection of an appropriate assurance provider and on good practice regarding the assurance process and statement elements.

Previous research on assurance statement quality typically draws on a combined list of assurance statement elements that are recommended by assurance standards (O'Dwyer & Owen, 2005; Perego & Kolk, 2012). This list serves as a benchmark to evaluate assurance statements through content analysis. Similarly, based on requirements and recommendations of ISAE 3000, AA1000AS, and GRI, we identify a benchmark list of assurance statement elements that guides the assessment of our main variable, *AI*, and assurance-related control variables in our model. Table 1 presents an overview of these statement elements by responsible party on the vertical axis and by research design on the horizontal axis.⁷

⁶ ISAE 3000 was recently revised. The new version will be effective for engagements that date the assurance statement on or after December 15, 2015.

⁷ The statement elements are described in greater detail in the subsequent sections of the paper.

Table 1. Overview of assurance statement elements

	Assurance intensity	Control variables	Complementary elements
Client	<ul style="list-style-type: none"> • <i>balance^a</i> • <i>level</i> • <i>scope</i> • <i>coverage</i> • <i>recommendations^a</i> 	<ul style="list-style-type: none"> • <i>addressees</i> 	<ul style="list-style-type: none"> • <i>responsibilities</i>
Provider		<ul style="list-style-type: none"> • <i>competence^b</i> • <i>independence^c</i> • <i>assuror</i> 	<ul style="list-style-type: none"> • <i>procedures</i> • <i>opinion</i> • <i>limitations</i> • <i>miscellaneous^d</i>

^a Item only mandatory/recommended according to AA1000AS/GRI.

^b Item only mandatory according to AA1000AS.

^c Item only mandatory according to AA1000AS/ISAE 3000.

^d Item includes representative, date, place, and standard.

In our research design, we distinguish among elements that describe the intensity of the assurance engagement, elements that refer to a statement's overall stakeholder orientation and to provider characteristics, and complementary elements. The first group constitutes elements that are used to construct our main variable of interest, namely, *AI*, and elements in the second group serve as control variables in our model. The elements in the third group complete a statement and either depend on elements in the aforementioned groups or refer to unrelated basic disclosures. All elements across the three groups are relevant for assessing the completeness of an individual assurance statement. The distinction between client and provider on the vertical axis indicates the party that is predominantly responsible for the definition of the assurance characteristics that ultimately translate into the respective statement elements. Although assurance statements are generally written by the assurance provider, the terms of the assurance engagement are negotiated between both parties prior to the engagement. The client, for instance, bears the costs of the assurance services and thus primarily determines the intensity-related elements that we use to proxy for *AI*. The assuror, however, "shall be satisfied that the engagement subject matter is appropriate" (AccountAbility, 2008, p. 13; similarly IFAC, 2008, p. 6). Ultimately, the assurance provider selects the appropriate assurance procedures and states his opinion.

Sample, data collection, and empirical model

Our initial sample consists of 195 companies located in France, Germany, Italy, Spain, Sweden, Switzerland, or the United Kingdom. To be included, a company must be listed in the Bloomberg European 500 index in January 2013, provide sufficient information on its CSP for the reporting year 2011, and belong to an industry group that contains at least five companies in total. Of these 195 companies, 68 percent seek external assurance. Further, of these 133 companies, eleven are excluded because the corresponding assurance statements could not be retrieved from corporate websites, a web search engine, or by direct request. Our final sample includes 122 European companies across 18 industry groups. The sample composition is depicted in Table 2.

Table 2. Sample composition by country⁸ and by industry

By country		n	By industry		n
1	France	26	1	chemicals	6
2	Germany	18	2	building materials, paper, steel	10
3	Italy	12	3	mining	8
4	Spain	19	4	advertising, entertainment, media	4
5	Sweden	5	5	telecommunication	9
6	Switzerland	11	6	computers and software	3
7	United Kingdom	31	7	car manufacturers	4
			8	retail	4
			9	food and beverages	5
			10	pharmaceuticals	6
			11	oil and gas	8
			12	banking and financial services	18
			13	insurance	4
			14	REITS	4
			15	aerospace and defense	5
			16	engineering and construction	9
			17	electricity	7
			18	gas, water, and electricity distribution	8
Total		122	Total		122

For each company, one of the authors manually recorded the information in the assurance statements between May and August 2014. All uncertainties were discussed between the authors and were resolved through consensus.

We use the following econometric model to test our hypotheses regarding whether companies on average capture sustainability assurance engagements in a beneficial or detrimental way:

$$\begin{aligned}
 \text{Equation (1): } AI = & \beta_0 + \beta_1 csp + \beta_2 addressee + \beta_3 independence + \beta_4 competence \\
 & + \sum_{i=5}^8 \beta_i assurator + \beta_9 fees + \beta_{10} size + \beta_{11} fin_perf + \beta_{12} leverage \\
 & + \sum_{j=13}^{29} \beta_j industry + \sum_{k=30}^{35} \beta_k country + \varepsilon.
 \end{aligned}$$

Assurance intensity

Our dependent variable *AI* captures the intensity of the assurance services requested by the client in a sustainability assurance engagement. *AI* is an index variable that covers five aspects of AI and that comprises the statement elements *balance*, *level*, *scope*, *coverage*, and *recommendations*. Each of these elements is measured as described below. Then, if necessary, the elements are proportionally rescaled on a [0, 1] interval and summed to the index variable *AI*.

⁸ No mandatory assurance requirements are in place in the countries considered for the reporting year 2011. In France, Grenelle II requires third-party assurance for companies exceeding certain size thresholds, but the stepwise phase-in process only began in 2012. Swedish companies are only obliged to seek external assurance if they are state owned. The 2011 Spanish Sustainable Economy Law requires companies to state whether their sustainability report has been assured, but there is no general obligation to seek external assurance.

All statement elements are mandatory, recommended, or allowed according to ISAE 3000, AA1000AS, and GRI, which ensures their measurability irrespective of the applied standard.

The first index element *balance* refers to an assurance engagement in which the assessor is requested to assess whether the sustainability report provides a balanced representation of the company's overall sustainability performance, i.e., "the report should reflect positive and negative aspects of the organization's performance to enable a reasoned assessment of overall performance" (GRI, 2013a, p. 17). Such a report avoids inflating a company's perceived sustainability performance by selective or overly positive disclosure. This aim is also reflected in the foundation principles of stakeholder inclusivity, materiality, and responsiveness within the AA1000AS. *Balance* is measured on a binary scale indicating whether the overall balance of the report and adequate stakeholder inclusion is assessed (equals "1") or not (equals "0").

The second *AI* index element *level* corresponds to the assurance engagement risk of having not detected any circumstances that may call into question the assessor's opinion on the objects being examined. According to AA1000AS, "the assurance provider achieves high assurance where sufficient evidence has been obtained to support their statement such that the risk of their conclusion being in error is very low but not zero". Similarly, "the assurance provider achieves moderate assurance where sufficient evidence has been obtained to support their statement such that the risk of their conclusion being in error is reduced but not reduced to very low but not zero" (AccountAbility, 2008, p. 11). Instead of "high" and "moderate", the ISAE 3000 proposes the terms "reasonable" and "limited". A reasonable/high assurance level is required for a positive assurance opinion, whereas a limited/low assurance level only allows for a negative opinion. A higher level of assurance is more costly because additional and more intensive assurance procedures are required to sufficiently reduce assurance engagement risk. *Level* is therefore measured on an ordinal scale, where "1" refers to a moderate assurance level, "2" to a mixed assurance level that depends on the different objects being assured, and "3" refers to a high assurance level. If the engagement is simply a review and thus if no assurance level is provided, the statement is coded "0".

Next, the index element *scope* refers to the question of which disclosures are covered by the assurance engagement. While the ISAE 3000 standard merely refers to the "subject matter", the AA1000AS standard distinguishes between "type 1" and "type 2" assurance depending on the scope of the engagement. According to the GRI, assurance engagements may include "policies outlined and data published" as well as "processes and systems related to sustainability" (GRI, 2013b, p. 11). Our index variable *scope* indicates whether the assurance covers numerical values *or* qualitative statements (equals "1"), numerical values *and* qualitative statements (equals "2"), or the full report (equals "3").

Similar to *scope*, the fourth *AI* index element *coverage* captures the number of core sustainability performance indicators that are covered by the assurance. Drawing on our measurement scheme for *csp* outlined in the next section, we measure *coverage* as the proportion of the number of assured core sustainability performance indicators relative to the number of reported core indicators. The measure thus ranges in the interval [0, 1]. A higher proportion is associated with more intensive assurance work.

Recommendations is the fifth and final *AI* index element, which is mandatory according to AA1000AS and voluntary according to ISAE 3000. This part of an assurance statement typically provides improvements to a company's sustainability reporting, systems, and processes. The preparation of recommendations by the assurator requires additional work, which bears higher cost and indicates a more intensive assurance engagement. Accordingly, the variable is coded "1" if recommendations are included and "0" otherwise. Table 3 summarizes the five elements that constitute our index variable *AI*.

Table 3. Overview of elements contained in independent index variable *AI*

<i>AI</i> element	Coding	Example / Computation
<i>balance</i>	(1) yes	"the [...] principles of [stakeholder] inclusivity, materiality and responsiveness have been used as criteria against which to evaluate the Report" (Barclays)
	(0) no	"our engagement is limited to [...] the 'cement' business line [...] we did not perform any assurance procedures on other information presented" (HeidelbergCement)
<i>level</i>	(3) high/reasonable	"to provide high level of assurance" (CaixaBank)
	(2) mixed	"moderate assurance" along with "the indicators identified by the symbol [...] reasonable assurance" (GDF Suez)
	(1) moderate/limited	"we have carried out a limited assurance engagement" (Snam)
	(0) none	no level is assigned, common terms are "review" (Holmen), "certification", and "verification" (both Credit Suisse)
<i>scope</i>	(3) full report	"assurance engagement on all parts of the [...] report" (Bayer)
	(2) numerical and qualitative	"assurance engagement on environmental performance information [...], including the explanatory notes" (Allianz)
	(1) numerical or qualitative	"perform [...] procedures [...] on the following environmental, safety, social [...] performance indicators" (Tullow Oil)
<i>coverage</i>	[0, 1] interval	assured sustainability performance indicators included in measurement scheme for <i>csp</i> / reported sustainability performance indicators included in measurement scheme for <i>csp</i>
<i>recommendations</i>	(1) yes	"we make the following recommendations for the further development" (Deutsche Telekom)
	(0) no	no recommendations are presented

Table 3 presents the elements that are used to proxy for *AI* along with examples from sustainability assurance statements and the corresponding coding categories.

Corporate sustainability performance and control variables

Our first independent and main variable of interest with respect to our research question is *csp*. We follow the approach developed by Hummel and Schlick (2014) to measuring CSP on the environmental and social dimensions of sustainability. Based on the GRI reporting guidelines, these authors derive eight essential performance indicators that can be applied to firms in different industry groups. The scheme comprises four environmental and four social performance indicators that are evaluated based on hand-recorded data. After the raw indicator data are winsorized at a 10 percent level and proportionally [0, 1] normalized on an industry group basis, the indicators are aggregated into the total sustainability performance score by calculating their mean. Thus, *csp* is a continuous variable that ranges in the interval between

zero (worst-in-industry-group performance for all indicators) and one (best-in-industry-group performance for all indicators). An overview of the measurement approach is provided in Table 4.

Table 4. Indicators included in the measurement scheme for *csp* (Hummel & Schlick, 2014)

Performance indicator	Measurement	Unit
<i>Environmental dimension</i>		
energy consumption	(direct + indirect energy consumption) / number of employees ^a	MWh / employee
water withdrawal	(total water withdrawal – cooling water) / number of employees ^a	m ³ / employee
greenhouse gas emissions	(GHG protocol scope 1 + scope 2 emissions) / number of employees ^a	t / employee
total weight of waste	total weight of waste / number of employees ^a	t / employee
<i>Social dimension</i>		
employee turnover	total number of employees ^a who leave / number of employees ^a * 100	percent
lost time incident rate	number of incidents resulting in lost time from work / (total hours worked / 200,000)	incidents / h
employee training	total training time / number of employees ^a	h / employee
share of women in the highest corporate bodies ^b	total number of women in the highest corporate bodies ^b / total number of members of the highest corporate bodies ^b * 100	percent

^a If available, full-time equivalents; headcount otherwise.

^b Management board and supervisory board for two-tier system countries; board of directors for one-tier system countries.

This approach to measuring *csp* is especially useful for our research setting because sustainability reports typically refer to both environmental and social dimensions. Firms seek to protect their perceived legitimacy on both dimensions by increasing their sustainability report's credibility through external assurance. Thus, the applied measurement scheme directly addresses the two dimensions and comprehensively covers sustainability performance within them. We expect to obtain a positive sign for *csp* if the detrimental form of MC prevails and a negative sign if the beneficial form dominates on average.

In addition to *csp*, we rely on the literature on sustainability assurance to select the control variables that are included in our model. First, we control for *addressee*, which is a binary variable indicating whether the assurance statement is explicitly addressed to a broad audience, such as all stakeholders (equals "1"), or not (equals "0"). Previous studies often regard a low degree of stakeholder inclusivity as an indicator of a detrimental form of MC (Cooper & Owen, 2007; Deegan et al., 2006; O'Dwyer & Owen, 2005). *Addressee* proxies for the intended audience of the assurance statement, which can affect accountability and transparency toward external stakeholders. We therefore expect *addressee* to have a positive effect on *AI*.

Next, we control for two provider characteristics, namely, *independence* and *competence*, which are essential requirements according to ISAE 3000, AA1000AS, and GRI. Findings from the financial audit literature suggest that limited auditor independence corresponds to lower audit quality, particularly a lack of objectivity in the audit engagement and obliging audit assessments (Blay & Geiger, 2013; DeAngelo, 1981). This finding is qualified by studies that report no relationship or even a negative relationship depending on market conditions and audit firm rotation (Daniels & Booker, 2011; Jamal & Sunder, 2011). Common proxies for auditor independence in the financial audit literature are often derived from audit fees (Blay & Geiger, 2013; Markelevich & Rosner, 2013). However, a similar approach is not feasible with respect to the measurement of assessor independence because companies do not separately disclose their expenses for sustainability assurance services. We therefore rely on four conditions that might indicate constraints on assessor independence.⁹ Our index variable *independence* indicates potentially threatened provider independence if at least one of these constraints is present (equals “0”) or none is present (equals “1”). An independent assurance provider can more easily stipulate assurance services of higher intensity for the self-serving purpose of increasing revenues. Consequently, we expect to find a positive effect of *independence* on our dependent variable *AI*.

The variable *competence* refers to both the familiarity of the assurance provider with a specific company and the provider’s general proficiency. *Competence* is an index variable that is calculated as the sum of a variable indicating whether the assurance provider has also assured the previous sustainability report (equals “1”) or not (equals “0”) and a variable indicating the completeness of the assurance statement. Completeness is measured as the proportion of information elements that are contained in the assurance statements relative to our benchmark list of all statement elements (as indicated in Table 1). The combined proxy for *competence* ranges in the interval between zero and two. We expect to find a positive relationship between *competence* and *AI* because a highly competent assurance provider has the necessary skills and experience to provide not only superficial but also intensive assurance services. In addition, such an assessor is more likely to have client-specific knowledge that helps her to stipulate more intensive assurance services during pre-engagement negotiations (O’Dwyer, 2011; O’Dwyer et al., 2011).

Previous research shows that both the assurance services and their descriptions in assurance statements differ depending on the type of assurance provider, particularly accountant assessors versus consultant assessors (Mock, Strohm, & Swartz, 2007; O’Dwyer & Owen, 2005). *Assessor* refers to four dummy variables indicating each of the Big 4 accounting firms and a pool of other assessors, such as small accounting firms, consultancy firms, or certification bodies.¹⁰

⁹ These conditions include additional sustainability-related consulting or assurance services, congruence between assessor and financial statement auditor, the characterization of the engagement as an iterative process, and a negative or missing declaration of independence (AccountAbility, 2008; IFAC, 2013).

¹⁰ This dummy structure also allows us to account for certain firms having more than one assurance provider at a given time.

In addition to the variables derived from the assurance statement, we control for a selection of firm-specific variables. Voluntary sustainability assurance engagements entail additional costs that depend on the intensity of the services requested. Clients are engaged in cost-benefit considerations, and providers may have to justify additional services based on these considerations (Kolk & Perego, 2010; Mock et al., 2007; O'Dwyer, 2011; Park & Brorson, 2005; Simnett et al., 2009). Because separate data on sustainability assurance fees are not available, we draw on the total amount of a firm's expenditures on audit, audit-related, and other services. Our variable *fees* refers to such expenditures divided by the number of employees and normalized on an industry group basis. We have no expectation regarding the sign, as higher total fees may either ease the justification for the relatively minor additional fees of more intensive sustainability assurance services or hamper the adoption of more intensive services, potentially because the overall fees are already high.

We draw on previous empirical studies and further control for a company's size, financial performance, leverage, industry group, and country of domicile (Cho et al., 2014; Kolk & Perego, 2010; Ruhnke & Gabriel, 2013; Simnett et al., 2009). Our variable *size* is measured as the decimal logarithm of the number of employees at the end of the fiscal year. Large firms may have greater negotiating power than smaller clients and may face additional legitimacy pressure because of their outstanding public visibility (Branco & Rodrigues, 2008; Dawkins & Fraas, 2011). However, we have no expectation regarding the sign, as the total effect is unclear and our sample primarily comprises large firms. The financial performance proxy *fin_perf* is measured as the free cash flow in millions of euros per employee at the 2011 fiscal year end. We expect a positive sign, as financially more sound companies may have greater capacity to request intensive assurance services and as providers may face less resistance when suggesting additional services. *Leverage*, which refers to the informational needs of a company's creditors, is measured as the average total assets divided by average total common equity. Because creditors of a highly leveraged firm may demand higher security regarding sustainability-related risks to assess the long-term success of a company, *leverage* may have a positive effect on our dependent variable *AI*. However, high leverage may also indicate financial difficulties or a shareholder structure dominated by myopic investors who demand cost cutting. Accordingly, we have no expectation regarding the sign for *leverage*. The series of dummy variables denoted by *country* is included to account for regulatory, cultural, and societal differences depending on companies' country of domicile. Especially under the detrimental MC hypothesis, firms regard sustainability assurance as an inconvenient compulsory exercise, for instance, to comply with common industry practice. Practices as well as accountability and transparency toward external stakeholders vary across industries. We account for this variability by also including *industry* group dummies. Table 5 provides an overview of all the model variables along with descriptions and data sources.

Table 5. Variables overview

Variable	Description	Source of data
<i>AI</i>	intensity of the assurance services, measured as an index variable as described in Table 3 and in the “Assurance intensity” section	assurance statement, sustainability disclosure
<i>csp</i>	CSP of the client, measured as described in Table 4 and in the “Corporate sustainability performance and control variables” section	sustainability disclosure
<i>addressee</i>	addressee of the assurance statement, measured as a binary variable indicating all stakeholders (1) or not (0)	assurance statement
<i>independence</i>	independence of the assurator, measured as a binary index variable indicating potentially threatened independence (0) or not (1)	assurance statement, annual report
<i>competence</i>	competence of the assurator, measured as an index variable ranging from zero to two and referring to the familiarity with a client and general proficiency	assurance statement
<i>assurator</i>	assurator dummy variables referring to Deloitte, Ernst & Young, KPMG, PricewaterhouseCoopers, others, and the reference category is PricewaterhouseCoopers	assurance statement
<i>fees</i>	total expenditures of the client on audit, audit-related, and other services divided by the number of employees and normalized on an industry group basis	annual report
<i>size</i>	firm size of the client, measured as the log of total employees at the end of the fiscal year	annual report, sustainability disclosure
<i>fin_perf</i>	financial performance of the client, measured as the free cash flow (cash flow from operating activities – total capital expenditures) in millions of euros per employee at the end of the fiscal year	Bloomberg database
<i>leverage</i>	financial leverage of the client (average total assets / average total common equity) during the fiscal year	Bloomberg database
<i>industry</i>	industry group dummy variables as reported in Table 2, and the reference category is banking and financial services	Bloomberg database, annual report
<i>country</i>	country of domicile dummy variables as reported in Table 2, and the reference category is the UK	Bloomberg database

RESULTS

Descriptive analyses

Panel A of Table 6 presents descriptive statistics on the five *AI* index variables *balance*, *level*, *scope*, *coverage*, and *recommendations*. The results for *balance* reveal that 37 percent of our sample firms not only seek veracity assurance but also mandate their assurance providers to perform more intensive work, i.e., to assess the overall balance of the contents and stakeholder considerations within their sustainability reports. Of our firms, 69 percent request a moderate *level* of assurance, 17 percent rely on mixed levels, and only 3 percent seek a high assurance level. The remaining 11 percent of assurance statements do not provide an assurance level, indicating a superficial assurance engagement of low intensity. This trend of most firms seeking a rather low level of AI also applies to the *scope* of assurance engagements. In 93 percent of all cases, the engagement is restricted to selected numerical and/or qualitative statements, whereas only 7 percent of the sample firms seek assurance of the full report. A more positive picture is revealed by the results for *coverage* and *recommendations*. Of the

core sustainability indicators covered by our measurement scheme for *csp*, on average, 80 percent are verified by an external assurator. Of the sample firms, 55 percent also request their assurers to provide *recommendations* on future improvements. One assurance statement explicitly contains all recommendations prepared by the assurator, whereas some inform the reader that additional recommendations were revealed only to the client. Overall, the constituents of *AI* present a mixed picture of the intensity of the assurance services requested by our sample firms.

Table 6. Descriptive statistics for model variables

<i>Panel A: Descriptive statistics AI elements (original values before rescaling)</i>						
	0	1	2	3	mean	sd
(1.1) balance	63%	37%	.	.	0.37	0.48
(1.2) level	11%	69%	17%	3%	1.13	0.63
(1.3) scope	.	31%	61%	7%	1.76	0.58
(1.4) coverage	0.80	0.26
(1.5) recommendations	45%	55%	.	.	0.55	0.50

<i>Panel B: Descriptive statistics regression variables</i>						
	n	min	max	mean	sd	
(1) AI	122	0.83	5.00	2.68	0.99	
(2) csp	122	0.06	0.93	0.55	0.18	
(3) addressee	122	0.00	1.00	0.13	0.34	
(4) independence	122	0.00	1.00	0.26	0.44	
(5) competence	122	0.70	2.00	1.68	0.40	
(6) fees	122	0.00	1.00	0.39	0.36	
(7) size	122	1.93	5.46	4.56	0.68	
(8) fin_perf	122	−1.97	1.17	0.02	0.31	
(9) leverage	122	−72.03	62.48	6.04	10.98	

Panel A in Table 6 presents descriptive statistics for the variables included in our index variable *AI*. The categories “0” to “3” refer to *balance* and *recommendations* being binary variables, *level* and *scope* being ordinal variables, and *coverage* being a continuous variable. Panel B shows *AI* on an aggregated level along with the other variables included in our regression analysis.

The previous mixed findings directly translate into a mean of 2.68 and standard deviation of 0.99 for our dependent variable *AI*, which is presented in Panel B of Table 6 along with descriptive statistics for our regression variables. Because some firms in our sample obtain the maximum value of 5.00, sustainability assurance does not appear to be a window-dressing activity in general. Our variable *csp* ranges between a minimum value of 0.06 and a maximum value of 0.93, with a mean of 0.55. These results are similar to those of Hummel and Schlick (2014) and do not indicate selection bias in our sample with respect to *csp*. Descriptive statistics on *addressee* reveal that the assurance statements are rarely addressed to external stakeholders, which is in line with previous studies (Cooper & Owen, 2007; Deegan et al., 2006; Manetti & Toccafondi, 2012; O’Dwyer & Owen, 2005). Regarding provider-specific determinants of *AI*, *independence* is rather low with a mean of 0.26. Low levels of provider independence reflect the nature of sustainability assurance engagements, i.e., providers are often engaged in additional consulting or auditing services and iterative interaction with their cli-

ents. Our broad measurement approach for independence is able to account for the disparities in the understanding of independence by the assurance providers in our sample. For instance, the assurance statement of one company does not disclose that the same company also uses the assurance provider as a financial auditor. Instead, it contains an extensive declaration of complete independence including a reference that no “prohibited services” are provided. In contrast to the results for independence, the mean of our proxy for assessor *competence* is 1.68, which is relatively high. The majority of our sample firms use the same assessor from the previous year, and on average, the statements come close to our completeness benchmark. Moreover, the descriptive statistics on *fees* reveal that the total costs due to any type of services by auditors/assessors are 444 euro per employee and year on average. Similarly, *size* (original values) shows that 73,430 employees work for our sample firms on average during 2011. *Fin_perf* and *leverage* exhibit substantial heterogeneity with respect to the financial performance and capital structure of our sample firms, which results from the variety of industry groups contained in our sample.

Table 7 presents the correlation matrix for our model variables. Our dependent variable *AI* is positively and significantly correlated with provider *independence*, *competence*, and clients’ *fin_perf*, which is fully in line with our expectations for these variables. The sign for *csp* is negative but nonsignificant. We are cautious in interpreting this result, as a correlation analysis cannot account for additional factors that may affect the relationship between *csp* and *AI*. Interestingly, *AI* is negatively and significantly correlated with *fees*, which reveals the difficulties that assessors face when promoting new sustainability assurance services in companies with substantial existing spending on traditional services. *Size* is negatively related to *addressee* and positively related to *competence*, which indicates that larger firms are less likely to address a broad audience and to choose more competent assessors. The negative correlation between *size* and *fees* might indicate large clients’ negotiating power that outweighs the increased overall assurance costs arising from their more complex international business. The positive correlation between *fin_perf* and *competence* may reflect the preferred selection of a more competent assessor if a company is financially sound. Intuitively, positive relationships also exist between *fin_perf* and both *size* and *leverage*. Overall, the correlation coefficients are clearly below 0.4, indicating no multicollinearity problems.

Table 7. Correlation statistics

	(1) <i>AI</i>	(2) <i>csp</i>	(3) <i>addressee</i>	(4) <i>independence</i>	(5) <i>competence</i>	(6) <i>fees</i>	(7) <i>size</i>	(8) <i>fin_perf</i>	(9) <i>leverage</i>
(1)	1.0000								
(2)	-0.1400 (0.1240)	1.0000							
(3)	0.0291 (0.7502)	0.0003 (0.9975)	1.0000						
(4)	0.2233 (0.0134)	-0.0218 (0.8116)	0.0443 (0.6277)	1.0000					
(5)	0.1737 (0.0557)	0.0473 (0.6050)	-0.0143 (0.8759)	0.0536 (0.5577)	1.0000				
(6)	-0.1924 (0.0338)	-0.0976 (0.2851)	0.0710 (0.4370)	0.0162 (0.8591)	0.0676 (0.4593)	1.0000			
(7)	0.0353 (0.6995)	0.0372 (0.6841)	-0.1736 (0.0558)	-0.1417 (0.1194)	0.1690 (0.0627)	-0.1519 (0.0950)	1.0000		
(8)	0.1546 (0.0892)	0.0981 (0.2825)	0.0592 (0.5175)	-0.0239 (0.7938)	0.2462 (0.0063)	-0.1068 (0.2418)	0.3357 (0.0002)	1.0000	
(9)	-0.0944 (0.3008)	-0.1165 (0.2013)	-0.0068 (0.9412)	-0.0571 (0.5324)	-0.0271 (0.7672)	-0.0673 (0.4611)	0.0000 (0.9998)	0.1838 (0.0427)	1.0000

Table 7 presents bivariate Pearson correlation coefficients and p-values (in parentheses) for a two-tailed test of statistical significance for the variables used in the regression analysis. Statistics are presented for the full sample of 122 firms.

Regression analysis

The results from a multiple regression analysis with *AI* as the dependent variable are presented in Table 8. We find a negative and highly significant ($b_{csp} = -1.12$, $p < 0.01$, two-tailed) relationship between our main variables of interest, *csp* and *AI*. This finding supports the *beneficial* MC hypothesis, which posits a negative relationship between *csp* and *AI*. On average, poor sustainability performers seek more intensive external assurance services to enhance their internal systems, processes, and CSP. These poor sustainability performers benefit the most from an intensive assurance engagement because they are more likely to face threatened legitimacy and to need to demonstrate the seriousness of their endeavors. Accordingly, superior sustainability performers on average demand less intensive assurance services owing to cost-benefit considerations. Because the sustainability performance of superior sustainability performers is already high and because they are thus most likely regarded as legitimate by external stakeholders, these high performers need to neither improve their systems, processes, and CSP nor enhance their external credibility. Moreover, highly intensive assurance is not currently common practice in most industries, which mitigates intra-industry pressure to catch up with industry peers. While our results generally do not exclude the possibility of a *detrimental* form of MC, we find significant evidence for the existence of a *beneficial* form of MC, indicating that firms' motives and market conditions in 2011 result in positive effects on transparency and accountability toward external stakeholders.

Table 8. Regression results: main model

	Expected sign detrimental MC	Expected sign beneficial MC	Estimates
<i>intercept</i>			3.4970*** (4.1154)
<i>csp</i>	+	–	–1.1197*** (–3.0612)
<i>addressee</i>	+	+	0.6690** (2.4589)
<i>independence</i>	+	+	0.3577** (2.2447)
<i>competence</i>	+	+	0.5792*** (3.4653)
<i>fees</i>	?	?	–0.4750** (–2.3853)
<i>size</i>	?	?	–0.1585 (–1.0075)
<i>fin_perf</i>	+	+	0.8498*** (3.2320)
<i>leverage</i>	?	?	–0.0220*** (–2.8816)
<i>assuror dummies</i>			included
<i>industry dummies</i>			included
<i>country dummies</i>			included
observations			122
adjusted R-squared			0.6206
F-statistic			6.6561***

Table 8 reports ordinary least squares coefficient estimates and t-statistics (in parentheses) based on Huber-White robust standard errors. *** and ** indicate statistical significance at the 1% and 5% levels (two-tailed). We estimate equation (1):

$$(1) AI = \beta_0 + \beta_1 csp + \beta_2 addressee + \beta_3 independence + \beta_4 competence + \sum_{i=5}^8 \beta_i assuror + \beta_9 fees + \beta_{10} size + \beta_{11} fin_perf + \beta_{12} leverage + \sum_{j=13}^{29} \beta_j industry + \sum_{k=30}^{35} \beta_k country + \varepsilon.$$

The findings regarding our control variables *addressee*, *independence*, *competence*, and *fin_perf* are in line with our expectations, revealing positive effects of the variables on *AI*. The positive sign on *addressee* implies that firms electing to directly address all stakeholders tend to seek more intensive assurance. A possible explanation for this finding is that an assurance statement being addressed to a broad audience would call the seriousness and credibility of the client firm into question if the corresponding engagement was merely superficial. The findings further show that the two provider characteristics *independence* and *competence* positively influence *AI*. This finding may indicate that independent and competent assurers that have acquired client-specific knowledge presumably possess increased negotiating power during the pre-engagement period and that they are thus more likely to convince the client to purchase more intensive assurance services (O'Dwyer, 2011; O'Dwyer et al., 2011). Surprisingly, the results show that *fees* is negatively related to *AI*. However, one must recall that this variable not only captures fees for sustainability assurance services but also includes a firm's

total expenses for any type of audit, audit-related, and other services. Park and Brorson (2005) report a rather small proportion of four to six percent for the fees for sustainability assurance services relative to the fees related to financial audits. Nevertheless, our results indicate that even such a small proportion is important for our sample firms, which is consistent with the findings of previous researchers (Kolk & Perego, 2010; Mock et al., 2007; O'Dwyer, 2011; Park & Brorson, 2005; Simnett et al., 2009). Presumably, owing to cost-benefit considerations, firms with higher overall expenses relative to their industry peers are less willing to spend further funds on voluntary sustainability assurance. The importance of cost-benefit considerations is also reflected by the positive impact of *fin_perf*. Given both effects, firms on average appear to evaluate the benefits of sustainability assurance engagements rather conservatively. This conservatism is relevant for assurance providers because they are more likely to negotiate more intensive assurance services if the client's overall fees are low and financial performance is high. Against this background, assurance fees appear to be an important instrument during the ongoing crowding-out period in the market for sustainability assurance. Further, the negative relationship between *leverage* and *AI* indicates that creditors may share the conservative view on the financial benefits of intensive sustainability assurance. However, the negative coefficient could also result from myopic shareholders who seek to increase their return on investment through higher leverage and cost savings. We find no impact of *size* on *AI*, indicating that public visibility, legitimacy pressure, and negotiating power may not exclusively depend on firm size. In addition, our sample primarily comprises large firms. Finally, consistent with previous findings, we find significant differences (not tabulated) depending on a firm's *assuror* (Mock et al., 2007; O'Dwyer & Owen, 2005), *industry* (Cho et al., 2014; Simnett et al., 2009), and *country* (Kolk & Perego, 2010; Ruhnke & Gabriel, 2013; Simnett et al., 2009).

Supplemental analyses

We perform a number of supplemental analyses to check the robustness of our results. First, we use log-transformed market capitalization to proxy for *size*, we use both return on assets and return on equity to proxy for *fin_perf*, and we winsorize our *csp* raw data at the 5 percent and 1 percent levels (instead of the 10 percent level). The results are similar to those for our main model (untabulated). Next, one may argue that our results are primarily driven by poor sustainability performers, whereas the nature of MC is different for superior sustainability performers. Thus, we divide our sample into poor and superior sustainability performers based on the median values within each industry group, and we separately re-estimate the regression for each sub-sample. Our results remain stable for both groups, revealing a highly significant negative relationship between *csp* and *AI* (untabulated), which indicates the existence of a *beneficial* form of MC within both sub-samples. Third, we separately assess the relationship between *csp* and *AI* for each dimension of sustainability. To do so, we replace our variable *csp* with a corporate environmental performance variable (*cenp*) and a corporate social performance variable (*csop*). These new variables are calculated as the arithmetic means of the winsorized and transformed performance indicators for the respective dimensions. The results are reported in Table 9.

Table 9. Regression results: environmental and social dimensions

	Expected sign detrimental MC	Expected sign beneficial MC	Estimates <i>cenp</i> model	Estimates <i>csop</i> model
<i>intercept</i>			3.1583*** (3.7126)	3.2516*** (3.7143)
<i>cenp</i>	+	−	−0.5546** (−2.4466)	
<i>csop</i>	+	−		−0.7916** (−2.5883)
<i>addressee</i>	+	+	0.8196*** (3.0359)	0.6514** (2.2139)
<i>independence</i>	+	+	0.3744** (2.3274)	0.3518** (2.1253)
<i>competence</i>	+	+	0.5617*** (3.3222)	0.4979*** (2.8667)
<i>fees</i>	?	?	−0.5561*** (−2.7470)	−0.2663 (−1.2949)
<i>size</i>	?	?	−0.1320 (−0.8303)	−0.1128 (−0.6926)
<i>fin_perf</i>	+	+	0.7167*** (2.7088)	0.9576*** (3.4826)
<i>leverage</i>	?	?	−0.0237*** (−3.0792)	−0.0205** (−2.5948)
<i>assuror dummies</i>			included	included
<i>industry dummies</i>			included	included
<i>country dummies</i>			included	included
observations			122	122
adjusted R-squared			0.6108	0.6090
F-statistic			6.4257***	6.3853***

Table 9 reports ordinary least squares coefficient estimates and t-statistics (in parentheses) based on Huber-White robust standard errors. *** and ** indicate statistical significance at the 1% and 5% levels (two-tailed). We estimate equations (2) and (3):

$$(2) AI = \beta_0 + \beta_1 cenp + \beta_2 addressee + \beta_3 independence + \beta_4 competence + \sum_{i=5}^8 \beta_i assuror + \beta_9 fees + \beta_{10} size + \beta_{11} fin_perf + \beta_{12} leverage + \sum_{j=13}^{29} \beta_j industry + \sum_{k=30}^{35} \beta_k country + \varepsilon$$

$$(3) AI = \beta_0 + \beta_1 csop + \beta_2 addressee + \beta_3 independence + \beta_4 competence + \sum_{i=5}^8 \beta_i assuror + \beta_9 fees + \beta_{10} size + \beta_{11} fin_perf + \beta_{12} leverage + \sum_{j=13}^{29} \beta_j industry + \sum_{k=30}^{35} \beta_k country + \varepsilon.$$

Although this approach limits the variation in our data, it allows us to separately test the applicability of our reasoning for each dimension. Both environmental performance and social performance appear to be important for our sample firms, as we obtain negative and significant signs for *cenp* and *csop*. These results are consistent with those of our main model, providing additional support for the existence of a *beneficial* form of MC. Finally, we investigate the problem of selection bias due to missing data for certain sustainability performance indicators. Similar to Hummel and Schlick (2014), we use the Monte Carlo simulation method developed by Metropolis and Ulam (1949) to simulate the outcome of our regression mod-

el if performance data were complete.¹¹ Every missing raw performance value is replaced with a simulated value that is randomly drawn from the continuous interval $[0.8 \times \text{minimum per industry}; 1.2 \times \text{maximum per industry}]$. The resulting complete performance dataset contains both true and simulated performance values. Then, we re-estimate our main regression model and obtain the estimates for the coefficient on *csp* along with the corresponding p-values. Figure 1 depicts a scatter plot of both values after the simulation is repeated 100,000 times with new randomly filled missing values in every run.

Figure 1: Results from Monte Carlo simulation of *csp* coefficient

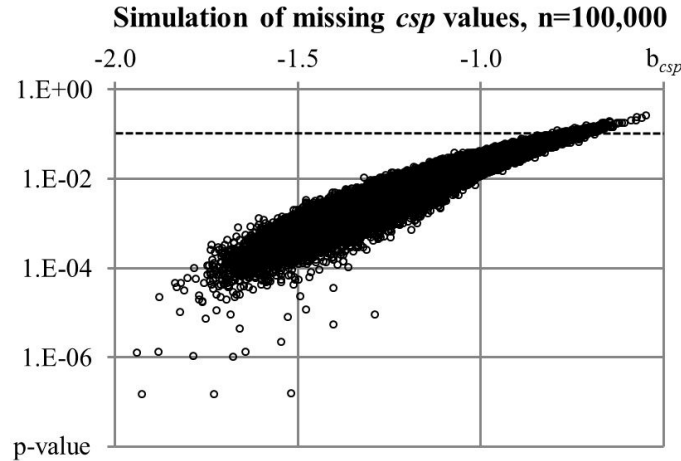


Figure 1 presents Monte Carlo simulation results to validate the robustness of our main regression model results to the simulation of missing data for certain sustainability performance indicators. Each mark represents a simulated regression coefficient along with the corresponding p-value. All marks below the broken line refer to statistically significant results at least at the 10% level (two-tailed).

The results regarding the existence of a beneficial form of MC are robust to the simulation of missing performance values, as the Monte Carlo simulation reveals generally significant negative values for the simulated coefficients on *csp*. Specifically, 54.29 percent of the coefficients are statistically significant at the 1 percent level, 42.83 percent are significant at the 5 percent level, 2.68 percent are significant at the 10 percent level, and 0.21 percent are nonsignificant.

CONCLUSIONS

This paper empirically investigates the nature of MC in sustainability assurance engagements. Previous research presents a predominantly critical picture of sustainability assurance as a process that is captured by the management of a company to advance its corporate image at the expense of accountability and transparency toward external stakeholders. Based on content analyses of assurance statements, such research often interprets the lack of transparency, accountability, and stakeholder inclusivity as evidence of the existence of a detrimental form of MC. Recent qualitative research based on interviews and case studies attenuates this critical picture and describes assurance providers as independent, competent, and highly commit-

¹¹ Firms must provide data on at least two performance indicators for both the environmental and social dimensions of sustainability (see Table 4) to qualify for our sample.

ted to accountability. Against the background of these ambiguous findings, we propose a broader and more neutral understanding of MC as *a client's influence on the process and outcome of assurance engagements*. Accordingly, we hypothesize that MC can affect assurance engagements in a *detrimental* or *beneficial* way with respect to transparency and accountability toward external stakeholders. We assess the nature of MC by investigating the relationship between the intensity of voluntarily requested sustainability assurance services and the client's CSP.

Drawing on a sample of 122 European companies, we observe a heterogeneous picture of AI, ranging from low-intensity to high-intensity assurance engagements. Although intensive sustainability assurance services are not currently common practice in most industries, the general perception of sustainability assurance as a detrimental, window-dressing activity is not supported by our findings. In particular, we find that AI is negatively related to CSP, indicating the existence of a beneficial form of MC. Poor sustainability performers on average demand more intensive assurance services to enhance their credibility toward external stakeholders, and the assurance process helps to improve their internal systems, processes, and CSP to close existing or impending legitimacy gaps. On the other hand, superior sustainability performers are already perceived as legitimate by their stakeholders, and they demand less-intensive assurance services owing to cost-benefit considerations.

Additionally, our study sheds light on other determinants of AI. Although the mechanisms could differ because of the voluntary nature of sustainability assurance engagements, our results corroborate related findings in the audit literature. An assessor's independence and competence appear to positively affect AI, as the position of the assessor during pre-engagement negotiations is strengthened and as client-specific knowledge helps to determine appropriate arguments to justify the costs of additional assurance services. Accordingly, relatively high client expenses on overall audit, assurance, and additional services are found to hamper AI, whereas sound financial performance positively affects the intensity of requested services. This picture is also consistent with current developments in the market for sustainability assurance, reflecting an ongoing crowding-out of smaller providers. Major accountancy firms benefit from their reputation as established financial auditors, and they may cross-sell sustainability assurance at competitive rates during this stage of the market. In line with previous research, we also find that the demand for assurance services differs depending on companies' industry and country of domicile.

Our findings and conclusions are moderated by certain limitations. First, the sample is restricted to large listed companies in 2011; thus, we cannot account for other companies that may receive less public attention and pursue different legitimization strategies. In addition, companies are excluded because of unavailable assurance statements or insufficient sustainability performance data. Second, data collection on AI relies on information contained in sustainability assurance statements, and this information is assumed to be true and complete. Although common assurance standards in 2011 provide guidance on assurance statement elements, latitude for interpretation and discretionary decisions by the assessors and their clients remains.

Finally, our investigation of the nature of MC in sustainability assurance engagements provides various opportunities for future research. In light of the revised ISAE 3000 standard that will be effective from the end of 2015 and in light of the ongoing developments in the market for sustainability assurance, researchers could analyze whether the major accountancy firms further strengthen their position and how assurance services are affected. In addition, our measure of AI could be applied to examine the effectiveness of forthcoming mandatory requirements. Such external shocks provide a setting for differences-in-differences analyses over time and for mandatory versus non-mandatory countries. Regarding the nature of MC, future research could investigate how regulation affects detrimental MC and whether beneficial MC diminishes because firms may refer to some legally required level of AI.

REFERENCES

- AccountAbility (2008). *AA1000 Assurance Standard 2008*. London: AccountAbility.
- Ameer, R., & Othman, R. (2012). Sustainability practices and corporate financial performance: A study based on the top global corporations. *Journal of Business Ethics*, 108(1), 61–79.
- Ball, A., Owen, D. L., & Gray, R. (2000). External transparency or internal capture? The role of third-party statements in adding value to corporate environmental reports. *Business Strategy and the Environment*, 9(1), 1–23.
- Blay, A. D., & Geiger, M. A. (2013). Auditor fees and auditor independence: Evidence from going concern reporting decisions. *Contemporary Accounting Research*, 30(2), 579–606.
- Branco, M. C., & Rodrigues, L. L. (2008). Factors influencing social responsibility disclosure by Portuguese companies. *Journal of Business Ethics*, 83(4), 685–701.
- Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? An empirical investigation of corporate environmental reputation. *Accounting, Organizations and Society*, 37(1), 14–25.
- Cho, C. H., Michelon, G., Patten, D. M., & Roberts, R. W. (2014). CSR report assurance in the USA: An empirical investigation of determinants and effects. *Sustainability Accounting, Management and Policy Journal*, 5(2), 130–148.
- Cho, C. H., & Patten, D. M. (2007). The role of environmental disclosures as tools of legitimacy: A research note. *Accounting, Organizations and Society*, 32(7–8), 639–647.
- Cooper, S. M., & Owen, D. L. (2007). Corporate social reporting and stakeholder accountability: The missing link. *Accounting, Organizations and Society*, 32(7–8), 649–667.
- Daniels, B. W., & Booker, Q. (2011). The effects of audit firm rotation on perceived auditor independence and audit quality. *Research in Accounting Regulation*, 23(1), 78–82.
- Darnall, N., Seol, I., & Sarkis, J. (2009). Perceived stakeholder influences and organizations' use of environmental audits. *Accounting, Organizations and Society*, 34(2), 170–187.

- Davis, K. (1973). The case for and against business assumption of social responsibilities. *Academy of Management Journal*, 16(2), 312–322.
- Dawkins, C., & Fraas, J. W. (2011). Coming clean: The impact of environmental performance and visibility on corporate climate change disclosure. *Journal of Business Ethics*, 100(2), 303–322.
- DeAngelo, L. E. (1981). Auditor independence, ‘low balling’, and disclosure regulation. *Journal of Accounting and Economics*, 3(2), 113–127.
- Deegan, C. (2002). Introduction: The legitimizing effect of social and environmental disclosures – A theoretical foundation. *Accounting, Auditing & Accountability Journal*, 15(3), 282–311.
- Deegan, C., Cooper, B. J., & Shelly, M. (2006). An investigation of TBL report assurance statements: UK and European evidence. *Managerial Auditing Journal*, 21(4), 329–371.
- Edgley, C. R., Jones, M. J., & Solomon, J. F. (2010). Stakeholder inclusivity in social and environmental report assurance. *Accounting, Auditing & Accountability Journal*, 23(4), 532–557.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston, MA: Cambridge University Press.
- GRI (2013a). *G4 reporting principles and standard disclosures*. The Netherlands: GRI.
- GRI (2013b). *The external assurance of sustainability reporting*. The Netherlands: GRI.
- Hodge, K., Subramaniam, N., & Stewart, J. (2009). Assurance of sustainability reports: Impact on report users’ confidence and perceptions of information credibility. *Australian Accounting Review*, 19(3), 178–194.
- Hummel, K., & Schlick, C. (2014). The relationship between sustainability performance and sustainability disclosure quality – Hard numbers beat smooth talk. *Working paper*, University of Zurich.
- IFAC (2008). *Assurance engagements other than audits or reviews of historical financial information* (2008 ed.). New York: IFAC.
- IFAC (2013). *Handbook of the code of ethics for professional accountants* (2013 ed.). New York: IFAC.
- Ihlen, Ø., & Roper, J. (2014). Corporate reports on sustainability and sustainable development: ‘We have arrived’. *Sustainable Development*, 22(1), 42–51.
- Jamal, K., & Sunder, S. (2011). Is mandated independence necessary for audit quality? *Accounting, Organizations and Society*, 36(4–5), 284–292.
- Kolk, A., & Perego, P. (2010). Determinants of the adoption of sustainability assurance statements: An international investigation. *Business Strategy and the Environment*, 19(3), 182–198.
- Kollman, K., & Prakash, A. (2001). Green by choice? Cross-national variations in firms’ responses to EMS-based environmental regimes. *World Politics*, 53(3), 399–430.

- KPMG (2013). *The KPMG survey of corporate responsibility reporting 2013*. Switzerland: KPMG.
- Manetti, G., & Toccafondi, S. (2012). The role of stakeholders in sustainability reporting assurance. *Journal of Business Ethics*, 107(3), 363–377.
- Markelevich, A., & Rosner, R. L. (2013). Auditor fees and fraud firms. *Contemporary Accounting Research*, 30(4), 1590–1625.
- Metropolis, N., & Ulam, S. (1949). The Monte Carlo method. *Journal of the American Statistical Association*, 44(247), 335–341.
- Mock, T. J., Strohm, C., & Swartz, K. M. (2007). An examination of worldwide assured sustainability reporting. *Australian Accounting Review*, 17(1), 67–77.
- O'Donovan, G. (2002). Environmental disclosures in the annual report: Extending the applicability and predictive power of legitimacy theory. *Accounting, Auditing & Accountability Journal*, 15(3), 344–371.
- O'Dwyer, B. (2011). The case of sustainability assurance: Constructing a new assurance service. *Contemporary Accounting Research*, 28(4), 1230–1266.
- O'Dwyer, B., & Owen, D. L. (2005). Assurance statement practice in environmental, social and sustainability reporting: A critical evaluation. *The British Accounting Review*, 37(2), 205–229.
- O'Dwyer, B., Owen, D. L., & Unerman, J. (2011). Seeking legitimacy for new assurance forms: The case of assurance on sustainability reporting. *Accounting, Organizations and Society*, 36(1), 31–52.
- Owen, D. L., Swift, T. A., Humphrey, C., & Bowerman, M. (2000). The new social audits: Accountability, managerial capture or the agenda of social champions? *The European Accounting Review*, 9(1), 81–98.
- Park, J., & Brorson, T. (2005). Experiences of and views on third-party assurance of corporate environmental and sustainability reports. *Journal of Cleaner Production*, 13(10–11), 1095–1106.
- Perego, P., & Kolk, A. (2012). Multinationals' accountability on sustainability: The evolution of third-party assurance of sustainability reports. *Journal of Business Ethics*, 110(2), 173–190.
- Pflugrath, G., Roebuck, P., & Simnett, R. (2011). Impact of assurance and assurer's professional affiliation on financial analysts' assessment of credibility of corporate social responsibility information. *Auditing: A Journal of Practice & Theory*, 30(3), 239–254.
- Power, M. (1991). Auditing and environmental expertise: Between protest and professionalisation. *Accounting, Auditing & Accountability Journal*, 4(3), 30–42.
- Power, M. (1996). Making things auditable. *Accounting, Organizations and Society*, 21(2–3), 289–315.

- Power, M. (1997). Expertise and the construction of relevance: Accountants and environmental audit. *Accounting, Organizations and Society*, 22(2), 123–146.
- Ruhnke, K., & Gabriel, A. (2013). Determinants of voluntary assurance on sustainability reports: An empirical analysis. *Journal of Business Economics*, 83(9), 1063–1091.
- Sethi, S. P. (1978). Advocacy advertising – The American experience. *California Management Review*, 21(1), 55–67.
- Simnett, R., Vanstraelen, A., & Chua, W. F. (2009). Assurance on sustainability reports: An international comparison. *The Accounting Review*, 84(3), 937–967.
- Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571–610.
- Wood, D. J. (1991). Corporate social performance revisited. *Academy of Management Review*, 16(4), 691–718.